

ISSN 2413 - 0133

Scientific journal

ИНФОРМАЦИОННЫЕ И МАТЕМАТИЧЕСКИЕ ТЕХНОЛОГИИ В НАУКЕ И УПРАВЛЕНИИ

№4(32)/2023

6+

**ИНФОРМАЦИОННЫЕ И МАТЕМАТИЧЕСКИЕ
ТЕХНОЛОГИИ В НАУКЕ И УПРАВЛЕНИИ**

Научный журнал

№ 4 (32)



EDITORIAL BOARD

Ablameyko S.V.
 Arshinskiy L.V.
 Berestneva O.G.
 Boukhanovsky A.V.
 Bychkov I.V.
 Woern H.
 Voevodin V.V.
 Wolfengagen V.E.
 Gornov A.Y.
 Gribova V.V.
 Groumpos P.
 Zorina T.G.
 Kazakov A.L.
 Kalimoldaev M.N.
 Karpenko A.P.
 Komendantova N.P.
 Kureichik V.V.
 Lis R.
 Massel L.V.
 Moskvichev V.V.
 Ovtcharova J.
 Popov G.T.
 Sidorov D.N.
 Smirnov S.V.
 Stennikov V.A.
 Stylios C.
 Taratukhin V.V.
 Khamisov O.V.
 Hodashinsky I.A.
 Chubarov L.B.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ ЖУРНАЛА

Абламейко С.В., академик НАН Беларуси, Минск, БГУ
 Аршинский Л.В., д.т.н., Иркутск, ИрГУПС
 Берестнева О.Г., д.т.н., Томск, ТПУ
 Бухановский А.В., д.т.н., Санкт-Петербург, НИУ ИТМО
 Бычков И.В., академик РАН, Иркутск, ИДСТУ СО РАН
 Вёрн Х., Германия, Karlsruhe Institute of Technology (KIT)
 Воеводин В.В., чл.-корр. РАН, Москва, НИВЦ МГУ
 Вольфенгаген В.Э., д.т.н., Москва, МИФИ
 Горнов А.Ю., д.т.н., Иркутск, ИДСТУ СО РАН
 Грибова В.В., чл.-корр. РАН, д.т.н., Владивосток, ИАПУ ДВО РАН
 Грумπος Π., Греция, University of Patras
 Зорина Т.Г., д.т.н., Республика Беларусь, Институт энергетики НАН Беларуси
 Казаков А.Л., д.ф.-м.н., профессор РАН, Иркутск, ИДСТУ СО РАН
 Калимолдаев М.Н., академик НАН РК, Республика Казахстан, ИИВТ
 Карпенко А.П., д.ф.-м.н., Москва, МГТУ им. Баумана
 Комендантова Н.П., PhD, Австрия, Лаксенбург, ПАСА
 Курейчик В.В., д.т.н., профессор ЮФУ, Таганрог
 Лис Р., Польша, Wrocław University of Science and Technology
 Массель Л.В., д.т.н., Иркутск, ИСЭМ СО РАН
 Москвичев В.В., д.т.н., Красноярск, ФИЦ ИВТ
 Овчарова Ж., Германия, Karlsruhe Institute of Technology (KIT)
 Попов Г.Т., Болгария, г. София, Технический университет
 Сидоров Д.Н., д.ф.-м.н., профессор РАН, Иркутск, ИСЭМ СО РАН
 Смирнов С.В., д.т.н., Самара, ИПУСС РАН
 Стенников В.А., академик РАН, Иркутск, ИСЭМ СО РАН
 Стилос Х., Греция, Technological Educational Institute of Epirus
 Таратухин В.В., Германия, ERCIS, University of Muenster
 Хамисов О.В., д.ф.-м.н., Иркутск, ИСЭМ СО РАН
 Ходашинский И.А., д.т.н., Томск, ТУСУР
 Чубаров Л.Б., д.ф.-м.н., Новосибирск, ФИЦ ИВТ

EXECUTIVE EDITORIAL

Chief Editor Massel L.V.
 Executive Editor
 Bakhvalova Z.A.
 Editor Gaskova D.A.
 Editor Massel A.G.
 Techn. Editor Popova M.D.
 Designer Pesterev D.V.

ИСПОЛНИТЕЛЬНАЯ РЕДАКЦИЯ ЖУРНАЛА

Главный редактор	Массель Л.В.	д.т.н., Иркутск, ИСЭМ СО РАН
Выпускающий редактор	Бахвалова З.А.	к.т.н., Иркутск, ИСЭМ СО РАН
Редактор	Гаськова Д.А.	к.т.н., Иркутск, ИСЭМ СО РАН
Редактор	Массель А.Г.	к.т.н., Иркутск, ИСЭМ СО РАН
Техн. редактор	Попова М.Д.	Иркутск, ИСЭМ СО РАН
Дизайнер	Пестерев Д.В.	Иркутск, ИСЭМ СО РАН

Учредитель: Федеральное государственное бюджетное учреждение науки
 Институт систем энергетики им. Л.А. Мелентьева Сибирского отделения Российской академии наук
 (ИСЭМ СО РАН)

Адрес учредителя, издателя и редакции

664033 г. Иркутск, ул. Лермонтова, 130
 Тел: (3952) 42-47-00 Факс: (3952) 42-67-96

Сайт журнала - <https://www.imt-journal.ru/>

Журнал включен в Российский индекс научного цитирования (РИНЦ). Номер контракта 202-04/2016.
 Журнал включен в перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней кандидата и доктора наук (по состоянию на 29.03.2022) по научным специальностям 1.2.2; 2.3.1; 2.3.3; 2.3.5; 2.4.5.
 Журнал зарегистрирован в Роскомнадзоре. Регистрационный номер ПИ № ФС 77 – 73539 от 31.08.2018.

Отпечатано в типографии «Дубль Принт»
 664046, г. Иркутск, ул. Волжская, 14

*Подписка на журнал доступна со второго полугодия 2021 г.
 Дата выхода 28.12.2023 г. Тираж 50 экз.

© Издательство ИСЭМ СО РАН

Цена свободная. (6+)

Методы искусственного интеллекта

- Дорогов А.Ю.** Быстрые преобразования и самоподобные нейронные сети глубокого обучения Часть 1. Стратифицированные модели самоподобных нейронных сетей и быстрых преобразований 5
- Чернавин П.Ф., Чернавин Н.П., Чернавин Ф.П.** Применение теории линейных неравенств в задачах машинного обучения 21
- Варламов О.О.** Создание Больших Знаний и расширение областей применения миварных технологий логического искусственного интеллекта 30
- Протасов В.И., Мирахмедов Р.О., Потапова З.Е., Чернова М.В.** Квазиоптимальное решение задачи коммивояжера с использованием метода эволюционного согласования 42
- Суров И.А.** Логика множеств и логика волн в когнитивно-поведенческом моделировании 51
- Алпатов А.В.** Применение машинного обучения для анализа образовательных результатов студентов вузов 67

Математическое моделирование в энергетике

- Репецкий О.В., Хоанг Д.К.** Математическое моделирование чувствительности собственных колебаний конструкций с учетом вращения и неравномерного нагрева 79
- Япарова Н.М., Капелюшин Ю.Е.** Идентификация внутреннего теплового источника и определение теплового состояния объекта по динамическим граничным измерениям 91
- Суджян А.М.** Исследование динамики реакторной установки ВВЭР-СКД при различных изменениях параметров питательной воды 104
- Еделев А.В., Карамов Д.Н., Башарина О.Ю.** Методика оценки живучести автономной микросети 117

Информационные технологии

- Берестнева О.Г., Лызин И.А., Аксенов С.В., Марухина О.В., Степаненко Н.П.** Разработка системы поддержки принятия врачебных решений на примере решения задачи выбора наилучшей траектории лечения для детей с эндокринопатиями 127
- Кузнецова А.С., Пушкарев А.А., Краснощеков К.В., Якубайлик О.Э., Ерунова М.Г.** Применение FABDEM и других современных цифровых моделей рельефа в системе аграрного мониторинга 139
- Гринберг Э.Я., Верзин Е.А.** Разработка и апробация методики картирования знаний наукоемкой организации 148
- Гасан В.С., Абасов Н.В., Осипчук Е.Н.** Мониторинг и накопление данных в информационно-прогностической системе ГеоГИПСАР 160
- Исаев С.В., Донцов Д.Ю.** Выделение групп используемых интернет-ресурсов для обнаружения внутренних источников киберугроз 170

Artificial Intelligence Methods

- Dorogov A.Yu.** Fast transformations and self-similar deep learning neural networks. Part 1. Stratified models of self-similar neural networks and fast transformations 5
- Chernavin P.F., Chernavin N.P., Chernavin F.P.** Application of the theory of linear inequalities in machine learning problems 21
- Varlamov O.O.** Creating Big Knowledge and expanding the applications of mivar technologies of logical artificial intelligence 30
- Protasov V.I., Mirakhmedov R.O., Potapova Z.E., Chernova M.V.** Quasi-optimal solution of the traveling salesman problem using the evolutionary matching method 42
- Surov I.A.** Logic of sets and logic of waves in cognitive-behavioral modeling 51
- Alpatov A.V.** Application of machine learning to analyze academic performance of university students 67

Mathematical modeling in the energy sector

- Repetckii O.V., Hoang D.C.** Mathematical modeling of the sensitivity of natural vibrations of structures, taking into account rotation and heating 79
- Yaparova N.M., Kapelyushin Yu.E.** Identification of an internal heat source and determining the thermal state of an object by dynamic boundary measurements 91
- Sujyan A.M.** Investigation of the dynamics of the VVER-SKD reactor plant with various changes in feed water parameters 104
- Edelev A.V., Karamov D.N., Basharina O.Yu.** Methodology for assessing the resilience of an autonomous microgrid 117

Information Technologies

- Berestneva O.G., Lyzin I.A., Aksenov S.V., Marukhina O.V., Stepanenko N.P.** Development of a medical decision support system on the example of solving the problem of choosing the best treatment trajectory for children with endocrinopathies 127
- Kuznetsova A.S., Pushkarev A.A., Krasnoshchekov K.V., Yakubailik O.E., Erunova M.G.** Application of FABDEM and other modern digital elevation models in the agricultural monitoring system 139
- Grinberg E.Ya., Verzin E.A.** Development and testing of the knowledge mapping methodology of a knowledge-intensive organization 148
- Gasan V.S., Abasov N.V., Osipchuk E.N.** Monitoring and accumulation of data in the GeoGIPSAR information-forecasting system 160
- Isaev S.V., Doncov D.Y.** Identification of groups of visited Internet resources for detection of internal cyberthreats source 170

Методы искусственного интеллекта

УДК 004.032.26

DOI: 10.25729/ESI.2023.32.4.001

Быстрые преобразования и самоподобные нейронные сети глубокого обучения. Часть 1. Стратифицированные модели самоподобных нейронных сетей и быстрых преобразований

Дорогов Александр Юрьевич

ПАО «Информационные телекоммуникационные технологии» («Интелтех»),
Санкт-Петербургский государственный электротехнический университет, Россия,
Санкт-Петербург, vaksa2006@yandex.ru

Аннотация. В работе показано, что в основе построения быстрых преобразований (подобных БПФ) лежат самоподобные структуры, которые в равной степени можно использовать и для построения быстрых нейронных сетей (БНС). Показано, что класс быстрых преобразований определяется системными инвариантами морфологического уровня и может быть описан, как морфогенез терминальных проекций нейронных модулей. Предложены лингвистические модели для описания морфологии, структуры и топологии регулярных самоподобных нейронных сетей. Модели легко обобщаются на многомерные варианты нейронных сетей данного типа. Благодаря своей структуре БНС обладают особыми алгоритмами обучения, которые принципиально отличаются от классического ErrorBackPropagation отсутствием механизма обратного распространения ошибки. Алгоритмы обучения основаны на предложенных в работе методах мультипликативной факторизации образов и быстрых преобразований. Разработанные алгоритмы завершаются за конечное число шагов с гарантируемой сходимостью (будет показано в части 2 статьи). Последовательное развитие концепции самоподобия приводит к разработке методов создания быстрых нейронных сетей с глубокой степенью обучения. Самоподобные нейронные сети обладают уникальной возможностью дообучения к новым данным без потери ранее приобретённых знаний. Показано, что БНС могут быть использованы для создания быстродействующей памяти образов с произвольным доступом и сложных устройств комбинационной логики. В работе представлены результаты исследований автора по следующим вопросам: биологические предпосылки самоподобия нейронных сетей; самоподобные многослойные структуры, морфогенез, стратификация модельных представлений; алгоритмы быстрых преобразований, быстрые нейронные сети (БНС), методы настройки; обучение БНС к эталонным функциям; пластичность БНС; пирамидальные нейронные сети глубокого обучения; многоканальные корреляторы; реализация памяти и комбинационной логики на пирамидальных структурах. Результаты исследований будут представлены статьями в трёх частях.

Ключевые слова: самоподобные структуры, быстрые преобразования, морфогенез, быстрые нейронные сети, пирамидальные структуры, память образов, комбинационная логика

Цитирование: Дорогов А.Ю. Быстрые преобразования и самоподобные нейронные сети глубокого обучения. Часть 1. Стратифицированные модели самоподобных нейронных сетей и быстрых преобразований / А.Ю. Дорогов // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 5-20. – DOI: 10.25729/ESI.2023.32.4.001.

Введение. Теория быстрых перестраиваемых спектральных преобразований появилась примерно в то же время, что и теория многослойных нейронных сетей. Обе теории развивались параллельно. Различная терминология, различные теоретические основы, различные области применения развели два направления достаточно далеко друг от друга, хотя общие черты – многослойность и перестраиваемость – были вполне очевидны.

Создание алгоритма быстрого преобразования Фурье (БПФ), безусловно, можно считать одним из выдающихся достижений второй половины 20 века. Время его рождения знаменательно совпало с начальным этапом развития вычислительной техники, когда быстродействие вычислительных машин было ещё крайне низким. Появление алгоритма БПФ стимулировало интерес и к другим видам спектральных преобразований. В задачах фильтрации, сжатия и выделения информативных признаков широкое применение нашли такие преобразования, как

Адамара-Уолша, Хаара, Виленкина-Кристенсона, Хартли, Наклонное, Вейвлет и другие, также обладающие быстрыми алгоритмами. Несмотря на отличия по видам функций, оказалось, что большинство алгоритмов быстрых преобразований имеют подобную структуру и отличаются друг от друга не более, чем значениями коэффициентов базовых операций. Осознание этого факта привело к идее построения обобщённых перестраиваемых спектральных преобразований, наделённых быстрым алгоритмом.

Возможность перестройки значений весовых коэффициентов и многослойная структура алгоритма роднят быстрые перестраиваемые преобразования с многослойными нейронными сетями прямого распространения. Первые подходы к обучению быстрых ортогональных преобразований были развиты в работах А.И. Солодовникова и его научной группы [1]. В то время подобный класс спектральных преобразований называли приспособленными быстрыми преобразованиями. Условие ортогональности влечёт за собой совпадение размерностей входного и выходного вектора для каждого слоя алгоритма. Если отказаться от этого условия, то снимается ограничение и на равенство размерностей, причём, как оказалось, структуру быстрого алгоритма при этом можно сохранить. Отсюда был прямой путь к построению быстрых нейронных сетей (БНС) [2, 3].

В рамках данной парадигмы быстрые перестраиваемые преобразования можно рассматривать, как многослойные нейронные сети, которые отличаются от последних линейными функциями активации и нулевыми смещениями в нейронах. По структурной организации БНС следует отнести к модульным нейронным сетям с инъективными связями. Роль модуля выполняют базовые операции малой размерности, например, для БПФ это базовая операция типа «бабочка». Дальнейшие исследования позволили доказать, что структура быстрых преобразований подчиняется системным инвариантам, основанным на принципе самоподобия.

Идея использования топологии быстрых преобразований БПФ для построения нейронных сетей многократно повторяется и в научных работах современных исследователей, но топология БПФ используется, как правило, только для проектирования нейросетевого модуля в составе свёрточной нейронной сети [4, 5]. Экспериментально показано, что это приводит к увеличению точности в задаче классификации и к увеличению быстродействия сетей. Для уменьшения числа вычислительных операций в свёрточных сетях, начиная с работы ЛеКуна [6], предлагается свёртку вычислять в частотной области на основе алгоритма БПФ. Эта же решение используется в ряде статей [7-9] для разных видов свёрточных сетей. Во всех выше перечисленных работах для построения алгоритма БПФ используется метод, предложенный Кули и Тьюки ещё в 60-х годах прошлого века для быстрого преобразования с размерностью, кратной степени двойки. Теоретически этот метод может быть использован и для других составных размерностей, но аналитика выражений оказывается весьма громоздкой. Упрощение способа построения преобразования является одной из целей настоящей работы. Далее будет показано, что это достигается за счёт стратификации модельных представлений быстрых алгоритмов.

Оказалось, что модульность и самоподобие характерны и для биологических нейронных сетей. В 20 веке идеи модульной самоподобной организации естественных нейронных сетей были последовательно высказаны нейробиологом Рафаэль Лоренте де Но (Lorente de No) [10], нейрофизиологами Вернон Бенджамин Маунткаслем (Vernon Benjamin Mountcastle) и Джералд Морис Эдельманом (Gerald Maurice Edelman) [11, 12], исследовавших строение новой коры головного мозга человека (неокортекса). Нейробиологи исторически называют нейронные модули нейронными колонками. Про новую кору известно, что вся она гораздо однороднее по своему строению, чем думали раньше; стремительное увеличение её в процессе филогенеза произошло путём умножения колонок, одинаковых в своей основе, а не путём развития

новых типов нейронов или разных способов внутренней организации. Внутренняя структура коры везде одинакова и состоит из повторяющихся многоклеточных единиц.

Лавинообразное увеличение неокортекса является важной чертой эволюции млекопитающих; степень этого увеличения отличает приматов от остальных млекопитающих, а человека – от остальных приматов. Если период появления ветви человека оценивают в 20 млн. лет, то развитие неокортекса – всего 2 млн. лет. Нейрофизиологи объясняют такое быстрое развитие многократным повторением структуры коры через однотипные многоклеточные единицы, фактически эволюционно обосновывая самоподобное строение неокортекса. Косвенным подтверждением этого уже в новом столетии явились результаты исследований [13], доказавшие, что все нейронные клетки в пределах одной мини-колонки связаны с одним и тем же рецептивным полем, а смежные мини-колонки могут быть связаны с разными рецептивными полями, что подтверждает принцип инъективной связи, являющийся необходимым условием самоподобия.

1. Стратификация модельных представлений. Математические модели служат средством изучения биологических нейронных сетей и проектирования искусственных нейронных сетей. От модели требуется, чтобы она была простой, но функционально достаточной. Основной проблемой математического моделирования является нахождение приемлемого компромисса между детализацией и простотой описания. Один из путей решения этой проблемы заключается в формировании иерархически вложенных семейств моделей, где каждый уровень иерархии соответствует уровню разумного абстрагирования свойств системы, что ведёт к упрощению каждой частной модели. Такое многоуровневое представление модели принято называть стратификацией [14], а каждый уровень модельного представления – стратой. Стратификация моделей позволяет разделить проектирование нейронной сети на относительно независимые этапы и использовать для каждого этапа специфичный математический аппарат.

Далее будет показано, что страты модельных представлений самоподобной нейронной сети упорядочены по степени абстракции. Высший уровень абстракции соответствует морфологическому представлению, а низший – параметрическому. В математической формулировке стратификация связана с выделением на каждом уровне иерархии эквивалентных отношений и переходом к факторным моделям, описывающим следующий уровень.

2. Морфологические модели модульных нейронных сетей. На морфологическом уровне игнорируется внутреннее строение модулей, значение имеет только их присутствие и наличие связей между ними. Характеристики связей также не определены. Ключевыми моментами концепции морфологического строения нейронных сетей являются многослойность, модульность и самоподобие. В этом разделе будут построены математические модели, связывающие эти понятия воедино.

2.1. Самоподобие и морфогенез. По определению, компактное топологическое пространство X самоподобно, если существует конечное множество S , индексирующее набор несюръективных отображений $\{f_s\}_{s \in S}$ для которых

$$X = \bigcup_{s \in S} f_s(X).$$

Пример. Компактом X является замкнутый отрезок $[0,1]$. Позиция точки на отрезке (кроме крайне правой точки) в десятичной системе выражается в виде $z = 0.z_{-1}z_{-2} \dots z_{-n} \dots$ в общем случае с бесконечным числом разрядов, где $z_i = 0, 1, 2, \dots, 9$.

Ограничим в позиционном представлении число разрядов уровнем n , и будем считать, что число $0.z_{-1}z_{-2} \dots z_{-n}$ соответствует отрезку длиной 10^{-n} на компакте X , а значение числа указывает начальную позицию этого отрезка на интервале $[0,1]$. В частности, для $n = 1$ число

$0.z_{-1}$ будет соответствовать отрезку длиной 0.1, а значение разряда z_{-1} указывает его позицию на X . Введём для этого уровня множество функций $f_{z_{-1}}(X)$, отображающих компакт X на отрезок длиной 0.1 в позиции $z = 0.z_{-1}$, очевидно, что

$$X = \bigcup_{z_{-1}} f_{z_{-1}}(X).$$

В этом примере символ объединения можно заменить прямой суммой, поскольку при разбиении отрезки не пересекаются. Таким образом, можно записать:

$$X = \bigoplus_{z_{-1}} f_{z_{-1}}(X).$$

Процесс можно итеративно продолжить, выбирая в качестве компактов отрезки уже построенного разбиения. Тогда для уровня n получим:

$$X = \bigoplus_{z_{-1} z_{-2}} \dots \bigoplus_{z_{-n}} f_{\langle z_{-1} z_{-2} \dots z_{-n} \rangle}(X).$$

В данном выражении кортеж $\langle z_{-1} z_{-2} \dots z_{-n} \rangle$ определяет многомерный индекс множества несюръективных отображений, разбивающих компакт X на части. Для позиционной системы счисления соответствие между значением числа и его поразрядным представлением $z \leftrightarrow \langle z_{-1} z_{-2} \dots z_{-n} \rangle$ является взаимно-однозначными.

Рассмотренное выше классическое понятие самоподобия ориентировано на обслуживание математических моделей фракталов, и не вполне подходит для представления самоподобных объектов не фрактальной природы, поэтому существует необходимость введения обобщающего определения, включающее в себя фрактал как частный случай. Если обратиться к биологии, то подходящим понятием может служить морфогенез живых систем. Под морфогенезом в биологии понимается процесс возникновения новых структур и изменения их формы в ходе индивидуального развития организмов. Не ставя перед собой задачу построения математической модели морфогенеза, введём рабочее определение, достаточное для задач морфологического синтеза самоподобных сетей.

Определение 1. Будем говорить, что на популяции индексированных объектов $X = \{A_z\}$ определён морфогенез, если для каждого объекта популяции найдётся точное соответствие «родитель-потомок», однозначно индексирующее дочерние объекты.

Например, пусть $X = X_0 = \bigoplus_{z_0} A_{z_0}$ – начальная индексированная популяция, тогда направленное развитие популяции по поколениям определяется последовательностью индексированных популяций вида:

$$X_0 = \bigoplus_{z_0} A_{z_0}, \quad X_1 = \bigoplus_{z_1} \bigoplus_{z_0} A_{\langle z_1 z_0 \rangle}, \dots, \quad X_{n-1} = \bigoplus_{z_{n-1}} \dots \bigoplus_{z_1} \bigoplus_{z_0} A_{\langle z_{n-1} \dots z_1 z_0 \rangle}, \dots$$

Если X – это компактное топологическое пространство, заданное отрезком единичной длины, и объектами популяций являются части отрезка, то все поколения объектов популяции совпадают с этим пространством, так что

$$X = X_0 = X_1 = \dots = X_{n-1} = \dots,$$

и если ещё потребовать конечности выбора правил индексации объектов популяции, то понятие морфогенеза трансформируется к определению самоподобия на компакте. Заметим теперь, что если ограничиться только условием конечности выбора правил индексации, то это приводит к самоподобным растущим популяциям, не связанным с фракталами. Если снять условие конечности выбора правил индексации, то процесс роста популяции вследствие морфогенеза в общем случае не будет самоподобным.

При неопределённости числа поколений затруднительно ответить на вопрос о конечности выбора правил индексации и самоподобия. В этом случае необходимо вводить дополнительные ограничения, выделяющие самоподобный процесс, например, если для всех поколений морфогенеза отображения «родитель-потомок» совпадают, то, очевидно, морфогенез тривиально самоподобен. Однако класс популяций тривиального самоподобного морфогенеза достаточно узкий, и для практических целей целесообразно его несколько расширить.

Определение 2. Морфогенез назовём *регулярным*, если в каждом поколении отображение «родитель-потомок» совпадает для всех объектов и однозначно определяется номером поколения.

Очевидно, что тривиальный самоподобный морфогенез является частным случаем регулярного, когда отображения «родитель-потомок» для всех поколений совпадают. Дальнейшее изложение будет в основном связано с регулярным морфогенезом. Проследивается явная автомодельность в описании регулярного и тривиального самоподобного морфогенеза, поэтому и для регулярного морфогенеза мы будем использовать термин самоподобный, делая необходимые пояснения в тех случаях, когда требуется выделить отличия.

2.2. Морфогенез многослойной самоподобной сети. Пусть многослойный сетевой граф Γ имеет n слоёв. Обозначим через $\{A_i^m\}$ – множество вершин в слое m , где $m = 0, 1, \dots, n-1$. Множество вершин входного (нулевого) слоя назовём *афферентом* сети и обозначим $Aff(\Gamma)$, а множество вершин конечного слоя назовём *эфферентом*¹ сети и обозначим $Eff(\Gamma)$. Пусть A_i^m – некоторая вершина сети в слое m . Назовём афферентом вершины (далее обозначается $Aff(A_i^m)$) подмножество вершин афферента сети, связанных дугами с вершиной A_i^m , таким образом, имеем $Aff(A_i^m) \subset Aff(\Gamma)$. Аналогично введём понятия эфферента вершины $Eff(A_i^m)$, как подмножество вершин эфферента сети, связанного дугами с вершиной A_i^m , таким образом, $Eff(A_i^m) \subset Eff(\Gamma)$. Афференты и эфференты вершин будем также называть *терминальными проекциями*.

Обозначим через $\Gamma^{-1}(A_i^m)$ рецепторную окрестность вершины A_i^m , а через $\Gamma(A_i^m)$ её аксоновую окрестность. Правило построения графа зададим следующими выражениями:

$$\begin{aligned} Aff(A_i^m) &= \bigoplus_{A_k^{m-1} \in \Gamma^{-1}(A_i^m)} Aff(A_k^{m-1}), \\ Eff(A_i^m) &= \bigoplus_{A_k^{m+1} \in \Gamma(A_i^m)} Eff(A_k^{m+1}). \end{aligned} \quad (1)$$

Эти правила были введены автором в 2004 году [15] и названы условиями слабой связанности, поскольку их реализация порождает слабосвязанные сети. Символ прямой суммы подчёркивает, что для любой вершины сети терминальные проекции вершин её окрестности не пересекаются.

В данном случае объектами популяции морфогенеза являются проекции вершин слоёв на терминальные множества. Номер слоя соответствует номеру поколения, причём для афферентов номер поколения возрастает от конечного слоя к начальному, а для эфферентов – от начального к конечному. Начальными популяциями при этом являются вершины терминальных слоёв. Условия слабой связанности справедливы как для регулярного, так и не регулярного морфогенеза.

¹ В биологии под афферентами понимают чувствительные нервные окончания (чувствительные рецепторы), которые расположены в органах и тканях и способны воспринимать разного рода раздражения. Эфферентами называют выходы нейронных подсистем, воздействующие на двигательные нервные волокна.

По отношению к проекциям на терминальное поле сети множество вершин слоя разбивается на афферентные и эфферентные классы. Афферентному классу принадлежат вершины нейронного слоя, имеющие общее афферентное множество, назовём эти классы афферентными доменами и обозначим $Dom_p(A_{z^m}^m) = X_{i^m}^m$, здесь $A_{z^m}^m$ - один из представителей домена, i^m - порядковый номер домена в слое m . Аналогично, отношение проекции вершин слоя m на терминальное эфферентное поле сети разбивает множество вершин конечного слоя на классы – эфферентные домены, которые обозначим $Dom_g(A_{z^m}^m) = Y_{j^m}^m$, где j^m - порядковый номер эфферентного домена в слое m .

Будем полагать, что n – слойный граф сети состоит из одной компоненты связности и удовлетворяет условиям (1). В этом случае на основании первого условия каждая вершина выходного слоя будет связана со всеми вершинами входного слоя, это означает, что в выходном слое существует всего один афферентный домен. Пусть этот домен имеет номер 0. При регулярном морфогенезе размеры всех рецепторных окрестностей вершин выходного слоя одинаковы, положим их равными p_{n-1} . Афференты вершин окрестностей вершины конечного слоя (т.е. слоя $n-1$) не пересекаются и поэтому множество вершин слоя $n-2$ распадается на домены, которые пронумеруем индексом i_{n-1} . Индекс будет изменяться в диапазоне от 0 до $p_{n-1}-1$. Таким образом, единственный домен конечного слоя в результате первого шага морфогенеза трансформировался в p_{n-1} доменов слоя $n-2$, которые обозначим $Dom_p^{n-2}\langle i_{n-1} \rangle$. Будем полагать, что каждая вершина слоя $n-2$ имеет рецепторную окрестность, состоящую из p_{n-2} вершин. В слое $n-3$ домены в результате морфогенеза получают номера, определяемые двойным индексом $Dom_p^{n-3}\langle i_{n-2}i_{n-1} \rangle$. Таким образом, достигнув начального слоя, получим, что нумерация доменов этого слоя будет определяться сложным индексом $Dom_p^0\langle i_1i_2 \dots i_{n-2}i_{n-1} \rangle$. Но в начальном слое каждый афферентный домен состоит только из одной вершины, поэтому построенный сложный индекс будет определять и номер вершины начального слоя. Афферентные домены промежуточных слоев индексируются выражением $Dom_p^m\langle i_{m+1}i_{m+2} \dots i_{n-2}i_{n-1} \rangle$.

Аналогично, каждая вершина входного слоя связана со всеми вершинами выходного слоя, это означает, что во входном слое существует всего один эфферентный домен. Для регулярного морфогенеза зададим размеры аксоновых окрестностей по слоям числами g_0, g_1, \dots, g_{n-2} и будем использовать индексы $j_0, j_1, \dots, j_{n-3}, j_{n-2}$ для нумераций доменов. Повторяя последовательность регулярных разбиений, получим, что для конечного слоя номер домена будет определяться сложным индексом $Dom_g^{n-1}\langle j_0j_1 \dots j_{n-3}j_{n-2} \rangle$. Но в конечном слое каждый эфферентный домен состоит только из одной вершины, поэтому построенный сложный индекс будет определять и номер вершины конечного слоя. Эфферентные домены промежуточных слоев индексируются выражением $Dom_g^m\langle j_0j_1 \dots j_{m-2}j_{m-1} \rangle$. Следующая теорема является обоснованием использования термина «слабосвязанная сеть».

Теорема. О параллельных путях. В слабосвязанных сетях отсутствуют параллельные пути.

Доказательство. Предположим противное. Пусть существуют два параллельных отрезка путей, которые начинаются в вершине A' и сходятся в вершине B , проходя через её окрестные вершины A_1 и A_2 . Так как A' – общая точка параллельных отрезков и она предшествует вершинам A_1 и A_2 , то афференты окрестных вершин A_1 и A_2 пересекаются так, что $Afr(A_1) \cap Afr(A_2) = Afr(A')$, но это противоречит условию слабой связности. Поскольку утверждение справедливо для любой вершины, включая терминальные, то отсюда следует невозможность параллельных путей для всей сети.

Теорема о морфологии слабосвязанной сети. В каждом слое слабосвязанной сети с одной компонентой связности афферентные и эфферентные домены попарно пересекаются во всех возможных сочетаниях точно по одной вершине, т.е. для каждой парной комбинации афферентного и эфферентного доменов существует единственная вершина $A_{z^m}^m$, такая, что

$$Dom_p(A_{z^m}^m) \cap Dom_g(A_{z^m}^m) = A_{z^m}^m.$$

Доказательство. Нужно доказать, что все парные сочетания разнотипных доменов имеют не пустое пересечение и что эти пересечения состоят только из одной вершины слоя.

Докажем первое. Предположим, что в слое m существует пара разнотипных доменов, которые не пересекаются по вершинам слоя. В этом случае вершины афферентного домена не связаны с частью вершин эфферента сети и поэтому в сети выделяется две правых конуса вершин, не связанных между собой во всех последующих слоях. С другой стороны, вершины эфферентного домена не связаны с частью вершин афферента сети и в сети выделяются два левых конуса, не связанных между собой в предшествующих слоях. Поскольку домены не пересекаются, то в этом случае в сети существует две не связанные между собой компоненты, что противоречит условию теоремы об однокомпонентности сети.

Докажем второе. Предположим, что домены пересекаются по двум вершинам A_1^m и A_2^m . Тогда эта пара вершин имеет общую афферентную окрестность, т.е. в слое $m-1$ существует, по крайней мере, одна вершина A_k^{m-1} , связанная с вершинами A_1^m и A_2^m , и её эфферент по условию морфогенеза (1), должен быть равен прямой сумме эфферентов вершин ее аксоновой окрестности. В эту прямую сумму входят и вершины A_1^m и A_2^m , но поскольку, по предположению, они принадлежат одному эфферентному домену слоя, то для них выполняется $Eff(A_1^m) \cap Eff(A_2^m) \neq \emptyset$ (см. рис. 1). Таким образом, мы пришли к нарушению условия морфогенеза (1) по эфферентам вершин. Аналогично это положение доказывается с использованием эфферентной окрестности пары вершин A_1^m и A_2^m .

Из этой теоремы следует, что для самоподобной сети существует взаимно-однозначное соответствие между номером вершины в пределах слоя и индексами пары пересекающихся разнотипных доменов. Для регулярной сети соответствие может быть задано кортежем, представляющим собой поразрядное представление числа в многоосновной системе счисления:

$$z^m = \langle j_0 j_1 \cdots j_{m-1} i_{m+1} i_{m+2} \cdots i_{n-1} \rangle. \tag{2}$$

В этом кортеже допустима любая перестановка индексов.

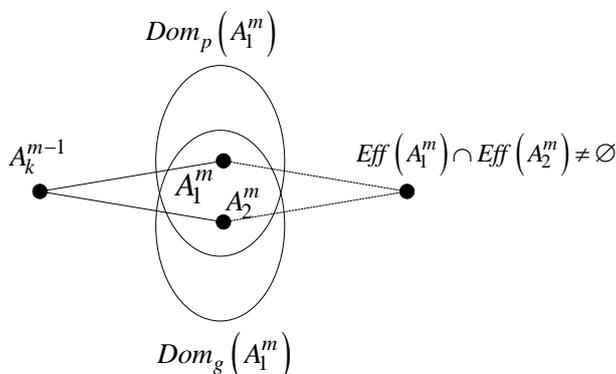


Рис. 1. Теорема о морфологии слабосвязанной сети.

3. Граф самоподобной сети. Из выражения (2), для слоя $m-1$ получим:

$$z^{m-1} = \langle j_0 j_1 \cdots j_{m-2} i_m i_{m+1} \cdots i_{n-1} \rangle.$$

Вершины смежных слоёв m и $m-1$ связаны дугой, если их афференты пересекаются. Афферент вершины z^{m-1} определяется разрядными числами кортежа $\langle i_m i_{m+1} \dots i_{n-1} \rangle$, а вершины z^m разрядными числами кортежа $\langle i_{m+1} i_{m+2} \dots i_{n-1} \rangle$, пересечение афферентов возможно только в том случае, когда одноименные разряды в приведённых кортежах будут совпадать. Такой же вывод можно сделать, используя условия пресечения эфферентов смежных слоёв m и $m+1$: вершины будут связаны дугой, если в кортежах $\langle j_0 j_1 \dots j_{m-1} \rangle$ и $\langle j_0 j_1 \dots j_{m-1} j_m \rangle$ одноименные разряды будут совпадать. Полученные правила позволяют достаточно просто выполнить построение графического образа самоподобной модульной сети. На рис. 2 показан пример построения четырёхслойной самоподобной сети для варианта, когда для всех m индексы принимают значения $i_m = \{0,1\}$ и $j_m = \{0,1\}$.

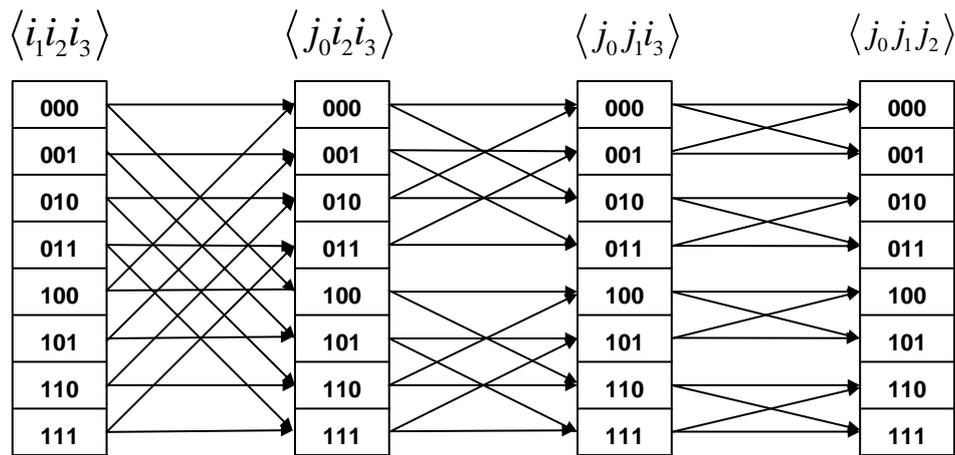


Рис. 2. Четырёхслойная самоподобная модульная сеть

Построенный граф полностью описывается лингвистическим предложением:

$$[\langle i_1 i_2 i_3 \rangle \langle j_0 i_2 i_3 \rangle \langle j_0 j_1 i_3 \rangle \langle j_0 j_1 j_2 \rangle].$$

Здесь каждое слово предложения представляет кортеж z^m номера вершины сети. В словах предложения допустима любая перестановка индексов.

4. Структурная модель модульной нейронной сети. Структурными характеристиками модульной нейронной сети являются размерности модулей по рецепторным и аксоновым полям и ранги межмодульных связей. Ограничим наше рассмотрение структурно-регулярной сетью, когда все модули в пределах слоя имеют одинаковые структурные характеристики, а все межмодульные связи имеют единичный ранг. Можно показать, что структурная регулярность влечёт за собой и морфологическую регулярность. Структурный синтез нерегулярных самоподобных сетей рассмотрен автором в работе [16].

Целесообразно упростить модель нейронной сети, полагая, что вся обработка сосредоточена в нейронных модулях, а связи осуществляют только передачу данных без внутренней обработки и искажений, т.е. осуществляют тождественное отображение с коэффициентом передачи равным единице. Такой случай характерен для алгоритмов быстрых преобразований типа БПФ.

Обозначим через p_m размерность рецепторных полей, а через g_m – размерности аксоновых полей нейронных модулей в слое m . Пример структурной модели модульной самоподобной нейронной сети с одноранговыми связями показан на рис. 3. Модули входного слоя имеют размерность $(3,2)$, а остальных слоёв – $(2,2)$. Веса всех дуг равны единицы и на рисунке не показаны. Структурная модель отличается от морфологической наличием весов вершин и дуг на графе модели.

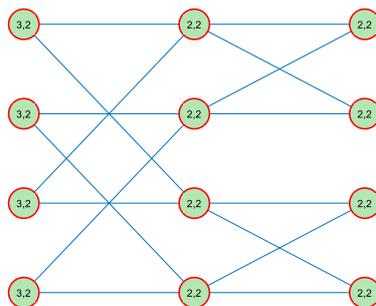


Рис. 3. Структурная модель регулярной самоподобной модульной нейронной сети с одноранговыми связями

5. Топологическая модель модульной нейронной сети. Входы и выходы структурной модели на рис. 3 являются трёх- и двух-координатными векторами без привязки к номеру отсчёта. Структурная модель непосредственно не переключается на алгоритм обработки в сети, её назначение – исследовать общие свойства модульной сети, связанные с вычислительной эффективностью и пластичностью. Для описания алгоритма необходимо построить топологическую модель. В топологической модели элементами рассмотрения являются физические контакты нейронных модулей; это либо входные рецепторы, либо выходные аксоны модуля.

Обозначим через u_m и v_m локальные номера рецепторов и аксонов для нейронного модуля в слое m . Диапазоны изменения локальных номеров определяются выражениями:

$$u_m = [0, 1, \dots, (p_m - 1)], \quad v_m = [0, 1, \dots, (g_m - 1)].$$

Позиционный номер рецептора в пределах нейронного слоя обозначим через U^m , а позиционный номер аксона через V^m . Совокупность взаимно-однозначных отображений вида:

$$\{u_m, z^m\} \rightarrow U^m, \quad \{v_m, z^m\} \rightarrow V^m$$

назовём *топологической моделью* нейронного слоя, здесь, как и прежде, z^m - номер модуля слоя m . Для регулярных нейронных сетей топологические отображения слоя могут быть выражены кортежами:

$$U^m = \langle \langle z^m \rangle \oplus u_m \rangle, \quad V^m = \langle \langle z^m \rangle \oplus v_m \rangle.$$

Символ \oplus в данном случае подчёркивает, что место размещения дополнительных разрядов u_m и v_m в кортеже $z^m = \langle j_0 j_1 \dots j_{m-1} i_{m+1} i_{m+2} \dots i_{n-1} \rangle$ может быть произвольным.

При одноранговых связях каждой дуге графа структурной модели однозначно ставится в соответствие дуга топологической модели, например, можно выбрать следующие взаимно-однозначные соответствия: $i_m \leftrightarrow u_m, \quad j_m \leftrightarrow v_m$, тогда возможен следующий вариант топологических отображений:

$$\begin{aligned} U^m &= \langle v_0 v_1 \dots v_{m-1} u_m u_{m+1} \dots u_{n-1} \rangle, \\ V^m &= \langle v_0 v_1 \dots v_{m-1} v_m u_{m+1} \dots u_{n-1} \rangle, \\ z^m &= \langle v_0 v_1 \dots v_{m-1} u_{m+1} \dots u_{n-1} \rangle. \end{aligned} \tag{3}$$

В особой ситуации находятся терминальные поля сети, поскольку для разрядных переменных u_0 и v_{n-1} нет соответствующих разрядов в кортеже z^m , но это не мешает построению кортежей топологических отображений слоя. Например, для структурной модели $[\langle i_1 i_2 i_3 \rangle \langle j_0 i_2 i_3 \rangle \langle j_0 j_1 i_3 \rangle \langle j_0 j_1 j_2 \rangle]$ можно построить следующие лингвистические предложения топологической модели:

Для рецепторных полей:

$$Rp = [\langle u_0 u_1 u_2 u_3 \rangle \langle v_0 u_1 u_2 u_3 \rangle \langle v_0 v_1 u_2 u_3 \rangle \langle v_0 v_1 v_2 u_3 \rangle].$$

Для аксоновых полей:

$$Ax = [\langle v_0 u_1 u_2 u_3 \rangle \langle v_0 v_1 u_2 u_3 \rangle \langle v_0 v_1 v_2 u_3 \rangle \langle v_0 v_1 v_2 v_3 \rangle].$$

Удобно эту модель представить одним лингвистическим предложением:

$$[\langle u_0 u_1 u_2 u_3 \rangle \langle v_0 u_1 u_2 u_3 \rangle \langle v_0 v_1 u_2 u_3 \rangle \langle v_0 v_1 v_2 u_3 \rangle \langle v_0 v_1 v_2 v_3 \rangle].$$

Первое слово в этом предложении соответствует рецепторному полю сети, а последнее – аксоновому. Граф топологической модели строится по тем же правилам, что и граф морфологической модели: дугами соединяются вершины, имеющие одинаковые значения разрядных переменных в смежных слоях. Граф топологической модели для данного примера показан на рис. 4.

Нетрудно убедиться, что данная модель соответствует графу быстрого преобразования Фурье (БПФ) в топологии Кули-Тьюки «с прореживанием по частоте» [17]. Другой вариант топологических отображений может быть задан в виде:

$$\begin{aligned} U^m &= \langle u_{n-1} u_{n-2} \cdots u_{m+1} u_m v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle, \\ V^m &= \langle u_{n-1} u_{n-2} \cdots u_{m+1} v_m v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle, \\ z^m &= \langle u_{n-1} u_{n-2} \cdots u_{m+1} v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle. \end{aligned} \quad (4)$$

Эта модель соответствует графу БПФ в топологии Кули-Тьюки «с прореживанием по времени». Граф данной топологической модели показан на рис. 5.

Из полученных результатов можно сделать вывод, что алгоритм быстрого преобразования Фурье является топологической реализацией самоподобной модульной сети, где модулями являются базовые операции типа «бабочка». Для быстрых нейронных сетей вместо базовой операции используется термин *нейронное ядро*.

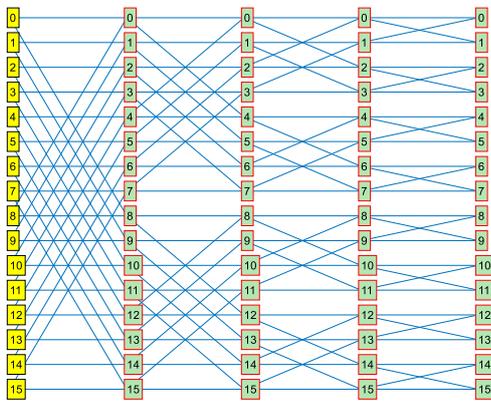


Рис. 4. Топологическая модель алгоритма БПФ «с прореживанием по частоте»

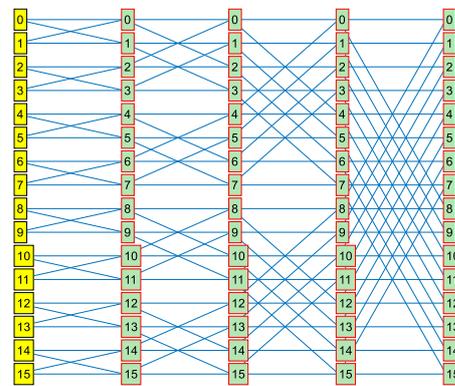


Рис. 5. Граф топологической модели алгоритма БПФ «с прореживанием по времени»

Нетрудно проверить, что для обеих выше приведённых топологий выполняется $V^{m-1} = U^m$, Из топологической модели (3) для терминальных слоёв получим:

$$U^0 = \langle u_{n-1} u_{n-2} \cdots u_1 u_0 \rangle, \quad V^{n-1} = \langle v_{n-1} v_{n-2} \cdots v_1 v_0 \rangle.$$

Если обозначить через N – размерность рецепторного поля, а через M – размерность аксонового поля сети, то из последних выражений непосредственно следует:

$$N = p_{n-1} \cdots p_1 p_0, \quad M = g_{n-1} \cdots g_1 g_0.$$

Таким образом, размерности терминальных полей сети определяются произведением размерностей нейронных модулей. На рис. 6 показан граф топологической модели (4) для структурной модели, показанной на рис. 3. Для наглядности на графе выделены базовые операции.

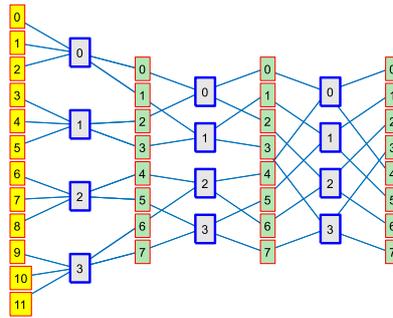


Рис. 6. Граф топологической модели с выделенными базовыми операциями

Рассмотрим ещё несколько типов регулярных топологий. На рис. 7 показан граф преобразования с топологией Гуда [18]. На графе явным образом выделены базовые операции. Работа Гуда по быстрому преобразованию Фурье появилась на семь лет раньше хорошо известной работ Кули и Тьюки [17], но прошла не замеченной.

Топологическая схема описывается моделью

$$U^m = \langle u_m u_{m+1} \dots u_{n-1} v_0 v_1 \dots v_{m-1} \rangle,$$

$$V^m = \langle u_{m+1} u_{m+2} \dots u_{n-1} v_0 v_1 \dots v_m \rangle,$$

$$i^m = \langle u_{m+1} u_{m+2} \dots u_{n-1} v_0 v_1 \dots v_{m-1} \rangle.$$

Особенность схемы Гуда состоит в том, что все слои графа имеют одинаковый топологический образ. Однако выходные операнды на каждой базовой операции не могут замещать входные отсчёты, поэтому требуется дополнительная память для хранения выходного вектора слоя. Рассмотренные выше топологические схемы Кули-Тьюки с «прореживанием по времени» и по «частоте» являются двойственными друг другу и могут быть получены зеркальным отражением относительно вертикальной оси. Алгоритмически переход к двойственной схеме выполняется заменой переменных в топологической модели по правилам: $U^m \leftrightarrow V^{n-1-m}$, $u_m \leftrightarrow v_{n-1-m}$. Двойственная топология Гуда описывается моделью:

$$U^m = \langle v_{m-1} v_{m-2} \dots v_0 u_{n-1} u_{n-2} \dots u_m \rangle,$$

$$V^m = \langle v_m v_{m-1} \dots v_0 u_{n-1} u_{n-2} \dots u_{m+1} \rangle,$$

$$i^m = \langle v_{m-1} \dots v_0 u_{n-1} u_{n-2} \dots u_{m+1} \rangle.$$

На рис. 8 показан граф быстрого преобразования с двойственной топологией Гуда.

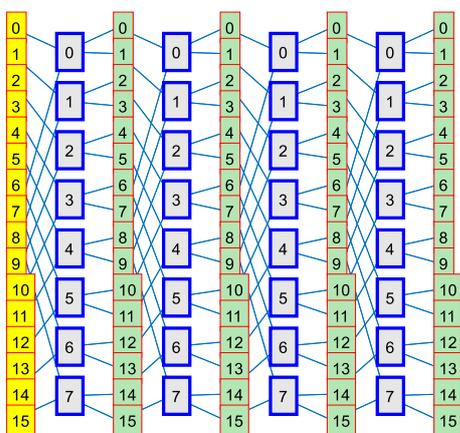


Рис. 7. Граф быстрого преобразования с топологией Гуда

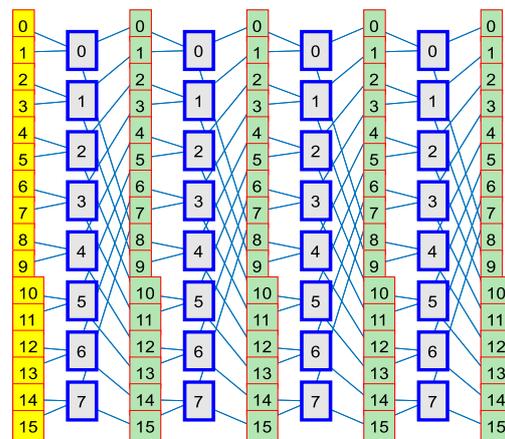


Рис. 8. Граф быстрого преобразования с двойственной топологией Гуда

Алгоритм БПФ для схем Кули-Тьюки по основанию 2 на выходе имеет спектр, гармоники которого имеют двоично-инверсную упорядоченность по частоте. Следующая топологическая схема позволяет получить на выходе спектр с естественным упорядочением по частоте при произвольных основаниях.

$$U^m = \langle u_m u_{m+1} \cdots u_{n-2} u_{n-1} v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle,$$

$$V^m = \langle u_{m+1} u_{m+2} \cdots u_{n-1} v_m v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle,$$

$$z^m = \langle u_{m+1} \cdots u_{n-2} u_{n-1} v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle.$$

На рис. 9 показан граф быстрого преобразования, построенный по данной топологии. Двойственная топологическая модель описывается выражениями:

$$U^m = \langle v_{m-1} v_{m-2} \cdots v_1 v_0 u_m u_{m+1} \cdots u_{n-2} u_{n-1} \rangle,$$

$$V^m = \langle v_m v_{m-1} v_{m-2} \cdots v_1 v_0 u_{m+1} \cdots u_{n-2} u_{n-1} \rangle,$$

$$z^m = \langle v_{m-1} v_{m-2} \cdots v_1 v_0 u_{m+1} \cdots u_{n-2} u_{n-1} \rangle.$$

Граф алгоритма БПФ с естественным упорядочением спектра по частотам и двойственной топологией показан на рис. 10.

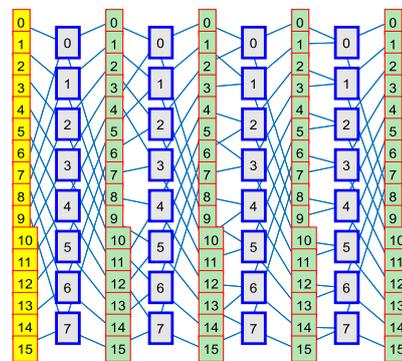
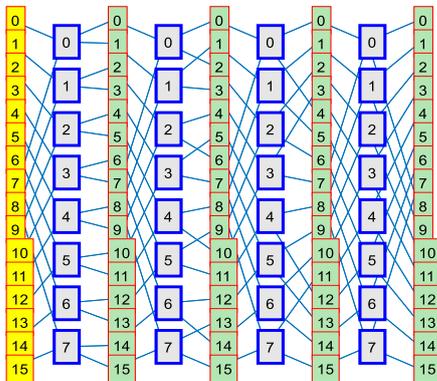


Рис. 9. Граф алгоритма БПФ с естественным упорядочением спектра по частотам следования

Рис. 10. Граф алгоритма БПФ с естественным упорядочением спектра по частотам следования и с двойственной топологией

Для быстрого преобразования Хаара используется следующая топологическая схема [3]:

$$U^m = \langle u_{n-1} u_{n-2} \cdots u_{m+1} u_m v_0 v_1 \cdots v_{m-2} v_{m-1} \rangle,$$

$$V^m = \langle u_{n-1} u_{n-2} \cdots u_{m+1} v_0 v_1 \cdots v_{m-2} v_{m-1} v_m \rangle,$$

$$z^m = \langle u_{n-1} u_{n-2} \cdots u_{m+1} v_0 v_1 \cdots v_{m-2} v_{m-1} \rangle.$$

Граф топологической схемы показан на рис. 11.

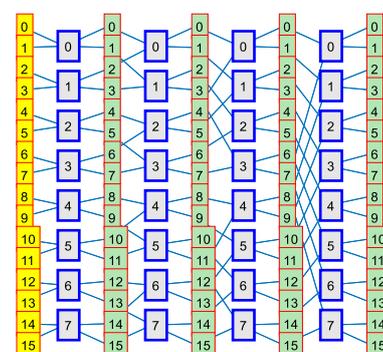
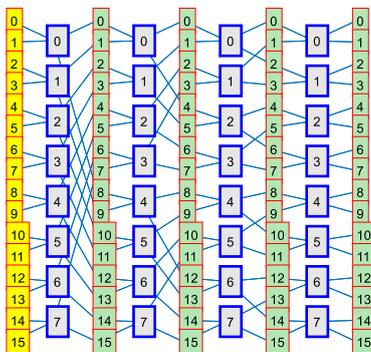


Рис. 11. Топологический граф преобразования быстрого преобразования Хаара

Рис. 12. Граф топологической схемы быстрого преобразования Хаара с двойственной топологией

Двойственна топологическая модель описывается выражением:

$$U^m = \langle v_0 v_1 \cdots v_{m-2} v_{m-1} u_{n-1} u_{n-2} \cdots u_{m+1} u_m \rangle,$$

$$V^m = \langle v_0 v_1 \cdots v_{m-2} v_{m-1} v_m u_{n-1} u_{n-2} \cdots u_{m+1} \rangle,$$

$$z^m = \langle v_0 v_1 \cdots v_{m-2} v_{m-1} u_{n-1} u_{n-2} \cdots u_{m+1} \rangle.$$

Топологический граф двойственной модели показан на рис. 12.

6. Параметрические модели быстрых преобразований. В графе быстрого алгоритма модулями являются базовые операции (нейронные ядра), представимые в виде матриц небольшой размерности. Для алгоритма быстрого преобразования в базовой операции z^m слоя m выполняется линейная обработка компонентов входного вектора слоя:

$$y_{z^m}^m(v_m) = \sum_{u_m} x_{z^m}^m(u_m) w_{z^m}^m(u_m, v_m),$$

где $x_{z^m}^m$ и $y_{z^m}^m$ – координаты входного и выходного векторов базовой операции; $w_{z^m}^m$ – матрица весов базовой операции. Для построения алгоритма преобразования необходимо перейти от локальных переменных ядра к глобальным переменным слоя. Этот переход реализуется на основе топологической модели. Параметрическое описание базовых операций вместе с топологической моделью образует параметрическую модель быстрого преобразования.

Заключение. В статье рассмотрен принцип построения четырехуровневой стратифицированной модели для самоподобных многослойных нейронных сетей с регулярной топологией. Стратификация позволяет исследовать систему на разных стадиях познания и описывать каждый уровень адекватными средствами. Стратифицирование можно рассматривать и как средство последовательного углубления представления о системе: при спуске вниз по иерархии страт система раскрывается в деталях; при подъёме на более абстрактные уровни яснее становится смысл и значение всей системы. В статье показано, что быстрые алгоритмы спектральных преобразований относятся к категории самоподобных нейронных сетей и могут быть представлены четырёхуровневой стратифицированной моделью.

Теоретическая основа быстрых алгоритмов долгое время базировалась на всевозможных теоремах факторизации, которые доказывали возможность разложения матрицы спектрального преобразования в произведение слабозаполненных матриц, где каждая матрица соответствует одному слою быстрого алгоритма [19, 20]. Это породило всплеск работ по теоремам факторизации. На этом пути исследователи столкнулись с тем обстоятельством, что существует большое множество различных матричных разложений для одного и того же спектрального преобразования. Когда число всевозможных теорем факторизации превысило десятки, стало понятным, что этот путь является тупиковым. Тем не менее, поток теорем не закончился и до сих пор. Проблема была чисто методическая и заключалась в смешивании понятий морфологии, структуры и топологии быстрого преобразования. Морфология является устойчивым системным инвариантом, свойственным всему классу быстрых алгоритмов, а топология – это не более чем допустимая реализация системного инварианта в связях между базовыми операциями. Каждая теорема факторизации соответствует одной из допустимых форм топологической реализации, а их число быстро растёт с ростом размерности преобразования. Автором было показано [21], что в основу теории быстрых преобразований следует положить именно системные инварианты, это устраняет необходимость изобретения новых теорем факторизации и позволяет предложить общий метод построения различных топологических реализаций быстрого преобразования при неизменной структуре. Предложенный метод существенно упрощает аналитику при построении быстрых преобразований с произвольными составными размерностями. Более того, выделение структурного и топологического уровней

даёт возможность решать новые задачи, связанные с оценкой параметрической и топологической пластичности перестраиваемых преобразований.

Список источников

1. Солодовников А.И. Основы теории и методы спектральной обработки информации / А.И. Солодовников, А.М. Спиваковский. – Л., 1986. – 272 с.
2. Дорогов А.Ю. Быстрые нейронные сети / А.Ю. Дорогов. – СПб.: Изд-во С.Петербург. ун-та, 2002. – 80 с.
3. Дорогов А.Ю. Теория и проектирование быстрых перестраиваемых преобразований и слабосвязанных нейронных сетей / А.Ю. Дорогов. – СПб.: «Политехника», 2014. – 328с.
4. Keivan Alizadeh Vahid, Anish Prabhu, Ali Farhadi, Mohammad Rastegari University of Washington butterfly transform: an efficient FFT based neural architecture design. CVPR-2020, Conference on computer vision and pattern recognition June 14-19, 2020, pp. 12024-12033.
5. Yingzhou Li, Xiuyuan Cheng, Jianfeng Lu Butterfly-Net: Optimal function representation based on convolutional neural networks, 2019, DOI: 10.48550/arXiv.1805.074512.
6. Mathieu M., Henaff M., LeCun Y., Fast training of convolutional networks through FFTs, 2013, DOI: 10.48550/arXiv.1312.5851.
7. Varsha Nair, Moitrayee Chatterjee, Neda Tavakoli, Akbar Siami Namin, Craig Snoeyink Fast fourier transformation for optimizing convolutional neural networks in object recognition, 2020, DOI: 10.48550/arXiv.2010.04257.
8. Shaohua Li, Kaiping Xue1, Bin Zhu, Chenkai Ding, Xindi Gao, David Wei, TaoWan. FALCON: A Fourier transform based approach for fast and secure convolutional Neural Network predictions. CVPR-2020, Conference on Computer Vision and Pattern Recognition June 14-19, 2020.
9. Chenlin Meng, Linqi Zhou, Kristy Choi, Tri Dao, Stefano Ermon. Butterfly Flow: building invertible layers with butterfly matrices. Proceedings of the 39-th International conference on machine learning, Baltimore, Maryland, USA, PMLR 162, 2022.
10. Lorente de Nó R. Cerebral Cortex: Architecture, intracortical connections, motor projections. Physiology of the nervous system. Oxford University Press, New York, 1949, pp. 288-312.
11. Эделмен Дж. Разумный мозг / Дж. Эделмен, В. Маунткасл. – М.: Мир, 1981. – 135.
12. Vernon B Mountcastle Sensory hand: neural mechanisms of somatic sensation. Cambridge, Harvard university press, 2005.
13. Jones E.G. Microcolumns in the cerebral cortex [Free PMC Article (full text)]. Proceedings of the National Academy of Sciences of the United States of America: journal, 2000, vol. 97, no. 10, pp. 5019-5021, PMID 10805761.
14. Волкова В.Н. Основы теории систем и системного анализа: Учеб. для студентов вузов / В.Н. Волкова, А.А. Денисов. – СПб.: Изд-во СПбГТУ, 1999. – 512 с.
15. Дорогов А.Ю. Быстрые нейронные сети: Проектирование, настройка, приложения. // Лекции по нейроинформатике Ч.1. В тр. школы-семинара «Современные проблемы нейроинформатики», науч.-техн. конф. «Нейроинформатика-2004» 28–30 января 2004 г. Москва. – М.: МИФИ, 2004, с. 69-135.
16. Дорогов А.Ю. Структурный синтез модульных слабосвязанных нейронных сетей. Часть 2 Ядерные нейронные сети / А.Ю. Дорогов // Кибернетика и системный анализ, 2001. – № 4. – с.13-20.
17. Good I.J. The Interaction algorithm and practical fourier analysis. Journal of Royal Statistical Soseity. Ser.B.- 1958, vol. 20, no.2, pp. 361–372.
18. Cooley J., Tukey J. An algorithm for the machine calculation of complex Fourier series. Math. Comput, 1965, vol. 19, pp. 297-301.
19. Рабинер Л. Теория и применение цифровой обработки сигналов / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978. – 848 с.
20. Дагман Э.Е. Быстрые дискретные ортогональные преобразования / Э.Е. Дагман, Г.А. Кухарев // Новосибирск: Наука. – 1983. – 228 с.
21. Дорогов А.Ю. Структурные и топологические инварианты быстрых перестраиваемых преобразований Сб. тр. VIII Науч.-техн. Всерос. научн. конф. «Нейроинформатика-2006» 24-27 января 2006г. – М.: МИФИ, 2006. – Ч.1. – С.39-50.

Дорогов Александр Юрьевич. Доктор технических наук, доцент, профессор кафедры «Автоматики и процессов управления» Санкт-Петербургского государственного электротехнического университета (СПбГЭТУ) «ЛЭТИ», главный научный сотрудник ПАО «Информационные телекоммуникационные технологии». Основные направления исследований автора связаны с интеллектуальным анализом данных, цифровой обработкой сигналов, проектированием быстрых преобразований и нейронных сетей быстрого обучения, разработкой аналитических платформ, моделированием радио-телекоммуникационных систем. AuthorID: 17611, SPIN 8645-5873, ORCID 0000-0002-7596-6761, vaksa2006@yandex.ru, г. Санкт-Петербург, ул. Попова, 5.

UDC 004.032.26

DOI: 10.25729/ESI.2023.32.4.001

Fast transformations and self-similar deep learning neural networks Part 1. Stratified models of self-similar neural networks and fast transformations

Alexander Yu. Dorogov

PJSC "Information Telecommunication Technologies" ("Inteltech"),

St. Petersburg State Electrotechnical University, Russia, St. Petersburg, *vaksa2006@yandex.ru*

Abstract. The paper shows that the construction of fast transformations (similar to FFT) is based on self-similar structures that can equally be used to build fast neural networks (FNN). It is shown that the class of fast transformations is determined by system invariants of the morphological level and can be described as the morphogenesis of terminal projections of neural modules. Linguistic models are proposed to describe the morphology, structure and topology of regular self-made neural networks. The models are easily generalized to multidimensional variants of neural networks of this class. Due to their structure, FNN have special learning algorithms that are fundamentally different from the classic ErrorBackPropagation by the absence of a mechanism for error back propagation. The learning algorithms are based on the methods of multiplicative factorization of images and fast transformations proposed in the work. The developed algorithms are completed in a finite number of steps with guaranteed convergence. The consistent development of the concept of self-similarity leads to the development of methods for creating fast neural networks with a deep degree of learning. Self-similar neural networks have a unique opportunity to learn to new data without losing previously acquired knowledge. It is shown that FNN can be used to create high-speed random access image memory and complex combinational logic devices. The paper presents the results of the author's research on the following issues: biological prerequisites for self-similarity of neural networks; self-similar multilayer structures, morphogenesis, stratification of model representations; algorithms for fast transformations, fast neural networks, tuning methods; training of FNN to reference functions; plasticity of FNN; pyramidal neural networks of deep learning; multi-channel correlators; implementation of memory and combinational logic on pyramidal structures. The research results will be presented in three parts of the article.

Keywords: self-similar structures, fast transformations, morphogenesis, fast neural networks, pyramidal structures, image memory, combinational logic

References

1. Solodovnikov A.I., Spivakovskij A.M. Osnovy teorii i metody spektral'noj obrabotki informacii [Fundamentals of the theory and methods of spectral information processing]. Leningrad, 1986, 272 p.
2. Dorogov A.Yu. Bystryye neyronnyye seti [Fast neural networks]. St. Petersburg Publishing House. Univ., 2002, 80 p.
3. Dorogov A.Yu. Teorija i proektirovanie bystryh perestraivaemyh preobrazovanij i slabosvjazannyh neyronnyh setej [Theory and design of fast tunable transformations and loosely coupled neural networks]. St. Petersburg, Politehnika [Politehnika], 2014, 328 p.
4. Keivan Alizadeh Vahid, Anish Prabhu, Ali Farhadi, Mohammad Rastegari University of Washington butterfly transform: an efficient FFT based neural architecture design. CVPR-2020, Conference on computer vision and pattern recognition June 14-19 2020, pp. 12024-12033.
5. Yingzhou Li, Xiuyuan Cheng, Jianfeng Lu Butterfly-Net: Optimal function representation based on convolutional neural networks, 2019, DOI: 10.48550/arXiv.1805.074512.
6. Mathieu M., Henaff M., LeCun Y., Fast training of convolutional networks through FFTs, 2013, DOI: 10.48550/arXiv.1312.5851.
7. Varsha Nair, Moitrayee Chatterjee, Neda Tavakoli, Akbar Siami Namin, Craig Snoeyink Fast fourier transformation for optimizing convolutional neural networks in object recognition, 2020, DOI: 10.48550/arXiv.2010.04257.
8. Shaohua Li, Kaiping Xue1, Bin Zhu, Chenkai Ding, Xindi Gao, David Wei, TaoWan. FALCON: A Fourier transform based approach for fast and secure convolutional Neural Network predictions. CVPR-2020, Conference on Computer Vision and Pattern Recognition June 14-19 2020.
9. Chenlin Meng, Linqi Zhou, Kristy Choi, TriDao, StefanoErmon. Butterfly Flow: building invertible layers with butterfly matrices. Proceedings of the 39-th International conference on machine learning, Baltimore, Maryland, USA, PMLR 162, 2022.

10. Lorente de Nó R. Cerebral Cortex: Architecture, intracortical connections, motor projections. Physiology of the nervous system. Oxford University Press, New York, 1949, pp. 288-312.
11. Gerald M. Edelman, Vernon B. Mountcastle Razumnyy mozg [Intelligent brain]. M.: Mir, 1981, 135p.
12. Vernon B Mountcastle Sensory hand: neural mechanisms of somatic sensation. Cambridge, Harvard university press, 2005.
13. Jones E.G. Microcolumns in the cerebral cortex [Free PMC Article (full text)]. Proceedings of the National Academy of Sciences of the United States of America: journal, 2000, vol. 97, no. 10, pp. 5019-5021, PMID 10805761.
14. Volkova V.N., Denisov A.A Osnovy teorii sistem i sistemnogo analiza: Ucheb. dlya studentov vuzov [Fundamentals of systems theory and system analysis: Textbook. for university students]. St. Petersburg, Publishing house of St. Petersburg State Technical University, 1999. 512 p.
15. Dorogov A.Yu. Bystryye neyronnyye seti: Proyektirovaniye, nastroyka, prilozheniya [Fast neural networks: Design, configuration, applications]. Lekcii po nejroinformatike Ch.1. V tr. shkoly-seminara "Sovremennyye problemy nejroinformatiki", nauch.-tehn. konf. "Nejroinformatika" [Proceedings of the school-seminar "Modern problems of neuroinformatics", scientific and technical conf. "Neuroinformatics"]. Moscow, MIFI, 2004, pp. 69-135.
16. Dorogov A.Y. Strukturnyy sintez modul'nykh slabosvyazannykh neyronnykh setey. Chast' 2 Yadernyye neyronnyye seti [Structural synthesis of modular weakly connected neural networks. II Nuclear neural networks]. Kibernetika i sistemnyy analiz [Cybernetics and system analysis], 2001, vol. 37, no. 4, pp. 470-477.
17. Cooley J., Tukey J. An algorithm for the machine calculation of complex Fourier series. Math. Comput, 1965, vol. 19, pp. 297-301.
18. Good I.J. The Interaction algorithm and practical fourier analysis. Journal of Royal Statistical Soseity. Ser.B.- 1958, vol. 20, no.2, pp. 361-372.
19. Lawrence R. Rabiner, Bernard Gold Teoriya i primeneniye tsifrovoy obrabotki signalov [Theory and application of digital signal processing]. M., Mir, 1978, 848 c.
20. Dagman Je. E., Kuharev G.A. Bystrye diskretnyye ortogonal'nyye preobrazovaniya [Fast discrete orthogonal transformations] Publ., Novosibirsk: Nauk :, 1983. 228 p.
21. Dorogov A.Yu. Strukturnyye i topologicheskiye invarianty bystrykh perestraivayemykh preobrazovaniy [Structural and topological invariants of fast tunable transformations]. Sb. tr. VIII Nauch.-tekhn. Vseros. nauchn. konf. "Neyroinformatika-2006" 24-27 yanvarya 2006g [Sb. tr. VIII Scientific and technical. All-Russian scientific conf. "Neuroinformatics-2006" January 24-27, 2006], M, MEPhI, 2006, Part 1, pp.39-50.

Dorogov Alexander Yurievich. Doctor of technical sciences, associate professor, professor of the department of Automation and Control Processes of St. Petersburg state electrotechnical university (SPbETU) "LETI", Chief researcher of PJSC "Information telecommunication technologies. The main directions of the author's research are related to data mining, digital signal processing, the design of fast transformations and neural networks of fast learning, the development of analytical platforms, modeling of radio and telecommunications systems. AuthorID: 17611, SPIN 8645-5873, ORCID 0000-0002-7596-6761, vaksa2006@yandex.ru , St. Petersburg, st. Popova, 5.

Статья поступила в редакцию 13.06.2023; одобрена после рецензирования 16.10.2023; принята к публикации 13.12.2023.

The article was submitted 06/13/2023; approved after reviewing 10/16/2023; accepted for publication 12/13/2023.

УДК 004.855.5

DOI:10.25729/ESI.2023.32.4.002

Применение теории линейных неравенств в задачах машинного обучения

Чернавин Павел Федорович, Чернавин Николай Павлович, Чернавин Федор Павлович

Уральский федеральный университет, Россия, Екатеринбург, *p.f.chernavin@urfu.ru*

Аннотация. По мнению авторов, результаты теории линейных неравенств надо более широко использовать в задачах машинного обучения (МО). Для исключения избыточных неравенств при построении ансамблей на основе линейных разделителей следует использовать теоремы о зависимых неравенствах следствиях. Для обобщения результатов различных исследований следует использовать модели поиска максимально совместных подсистем. Понятие “максимально совместные подсистемы” следует расширить и распространить его на комитеты единогласия. Для решения задач классификации можно эффективно использовать метод выпуклых оболочек, а для определения экстремальных точек выпуклых оболочек – использовать результаты из теории альтернативных систем. В статье приведены необходимые для этого сведения из теории линейных неравенств, математические модели на их основе и листинг программ для компьютерной реализации математических моделей.

Ключевые слова: машинное обучение, линейные неравенства, максимально совместные подсистемы, математическое программирование, классификация, медицинская диагностика

Цитирование: Чернавин П.Ф. Применение теории линейных неравенств в задачах машинного обучения П.Ф. Чернавин, Н.П. Чернавин, Ф.П. Чернавин // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 21-29. – DOI: 10.25729/ESI.2023.32.4.002.

Введение. Развитием теории линейных неравенств занимались многие известные математики: Минковский Г., Фаркаш Б., Черников С.Н. [1], Еремин И. И. [2]. Ряд методов машинного обучения (МО) был создан на основе этой теории. Например, решающее правило (РП) метода линейного разделения множеств – это обычное линейное неравенство, а суть РП состоит в том, что объекты разных классов находятся по разные стороны разделяющей гиперплоскости, и если в уравнение гиперплоскости подставить входные параметры объекта наблюдения, то для объектов одного класса значение выходного параметра будет строго больше нуля, а для другого меньше или равно. Данное РП хорошо воспринимается практическими специалистами и имеет хорошо понимаемую геометрическую интерпретацию, от которой в большинстве случаев легко перейти к содержательной интерпретации и практическим рекомендациям.

Если на основе одного неравенства не удастся построить РП с желаемыми качественными показателями (Accuracy, Precision, AUC ROC и т.п), то обычно на основе нескольких неравенств удастся создать ансамбль, удовлетворяющий этим показателям. Мы предпочитаем использовать синонимы слова “ансамбль” – комитет или комитетные конструкции, так как, на наш взгляд, это название более точно отражает логику такого подхода к решению задач классификации. Комитеты могут иметь различную логику и поэтому в зависимости от нее имеют названия: комитет единогласия (КЕ), комитет большинства (КБ) и комитет старшинства (КС). Наиболее полное развитие метод комитетов получил в научной школе Мазурова В. Д. [3, 4]. В рамках этой школы доказано, что КБ и КС существуют всегда, поэтому этими комитетами теоретически можно решить любую задачу классификации. В данном случае мы просто хотим сказать, что этот подход является универсальным, но не единственным, и качество РП на его основе, конечно, всегда нужно сравнивать с качественными характеристиками РП на основе других методов МО.

Все комитетные конструкции имеют хорошую и понятную геометрическую интерпретацию. Различные виды графических изображений комитетных конструкций и практическое применение метода комитетов мы приводим в [5, 6]. Отметим также, что множества можно разделять и нелинейными функциями (желательно сепарабельными), так

как за счет перехода в другое пространство признаков разделение множеств можно свести к разделению линейными функциями [5].

Тем не менее, после всего сказанного, мы хотим добавить, что, на наш взгляд, многие достижения теории линейных неравенств слабо используются в задачах исследования операций и МО, или не используются совсем. В первую очередь это касается анализа полученных РП и выдачи рекомендаций на их основе. В большинстве случаев, для выдачи практических рекомендаций требуется решить другие задачи на основе полученных РП. Такие задачи методологически более правильно относить к задачам исследования операций, и в их решении теория линейных неравенств может быть хорошим помощником. В данной статье мы будем ссылаться в основном на примеры из медицинской диагностики так как они содержательно наиболее воспринимаемы. При этом отметим, что аналогичным образом мы решали задачи и в других предметных областях (металлургия, анализ финансовых рынков, банковский скоринг).

1. Модель поиска неравенств следствий. Проверку неравенства на следствие системы неравенств будем делать, исходя из теоремы Минковского-Фаркаша в изложении Черникова и Еремина [1,2]. В наиболее воспринимаемой форме данная теорема, на наш взгляд, изложена в монографии Зоркальцева В.И. [6].

Линейное неравенство

$$C^T x \geq d \quad (1)$$

назовем следствием системы линейных неравенств

$$Ax \geq b, \quad (2)$$

если любой вектор $x \in R^n$, удовлетворяющий системе (2), удовлетворяет условию (1). Пусть система (2) совместна. Неравенство (1) будет следствием системы (2), если при некотором

$$u \in R^n \quad A^T u = c, \quad b^T u \geq d, \quad u > 0 \quad (3)$$

Геометрически данная теорема означает, что вектор градиента неравенства, проверяемого на следствие, должен быть в конусе векторов системы неравенств.

Можно сказать, что любой комитет из линейных разделителей – это просто некоторая система неравенств, связанная с определенной логикой для выработки РП. Рассмотрим ситуацию, довольно часто встречающуюся в медицинской диагностике. Для конкретной нозологии построено решающее правило в виде КС. В большинстве случаев это означает, что множество наблюдений разбивается на несколько областей, так, что в одних сконцентрированы здоровые, а в других больные. Несмотря на то, что в большинстве такого рода задач рассматривается количество входных признаков значительно более двух, без потери общности, графические пояснения мы можем дать только на плоскости. Графическая интерпретация данной ситуации приведена на рисунке 1.

На данном рисунке голубые точки соответствуют здоровым, а красные больным. Из рисунка следует, что наблюдения, соответствующие здоровым, сконцентрированы в двух не пересекающихся областях. Для выдачи конкретных рекомендаций требуется определить, какими неравенствами ограничена каждая из областей. Из приведенного рисунка видно, что любое решение системы неравенств 1, 2, 3 будет решением и для неравенства 4, то есть данное неравенство является следствием системы 1, 2, 3 и при ее описании может быть исключено из рассмотрения. В свою очередь, неравенство 2 является следствием системы 1, 3, 4. В двухмерном пространстве, при наличии рисунка, неравенства следствия можно определить визуально. В многомерном пространстве визуально определить неравенство следствие нельзя, но это можно сделать на основе математической модели.

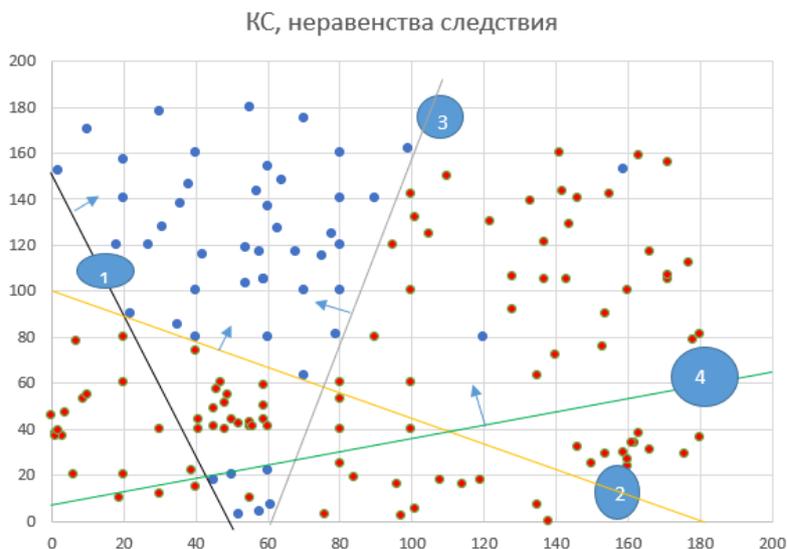


Рисунок 1. Графическая интерпретация комитета старшинства

В условиях теоремы сказано, что система должна быть совместной. С практической точки зрения это не совсем удобно, поэтому приведем математическую модель для определения неравенств следствий в пространстве любой размерности в кодах IBM ILOG CPLEX с учетом того, что система может быть и несовместной.

Листинг 1. Модель определения неравенств следствий

```

int n =...; //число входных признаков
range i = 1..k; // индекс входного признака
int m =...; //число неравенств в системе
range j = 1..m; // индекс неравенства системы
float A[j][i]=...; //коэффициенты системы неравенств
float B[j]=...; //правые части системы неравенств
float C[i]=...; // коэффициенты неравенства проверяемого на
следствие
float D=...; //правая часть неравенства проверяемого на следствие

/* искомые переменные */
dvar float+ u[j] ; //коэффициенты неравенств в линейной комбинации
dvar float+ v[i] ; //невязка коэффициентов линейной комбинации
dvar float+ w[i] ; //невязка коэффициентов линейной комбинации
dvar float+ svw ; //сумма невязок коэффициентов линейной комбинации
dvar float+ vd ; //запас по свободному члену неравенства следствия
dvar float+ wd ; //невязка правых частей линейной комбинации правых
частей

minimize svw;

subject to {

forall( i in i) sum (j in 1..m) A[j][i]*u[j]==C[i]+v[i]-w[i];
svw==sum(i in i) (v[i]+w[i]);
sum (j in 1..m) B[j]*u[j]==D-wd+vd;
forall( j in j) u[j] >=0;};
    
```

Если значение целевой функции (svw) равно нулю, то градиент неравенства, проверяемого на следствие, находится в конусе градиентов системы неравенств, но следствием оно является только в случае, если $vd > wd$, то есть неравенство, проверяемое на следствие, не пересекает систему неравенств.

Описанная выше схема: построение комитетной конструкции, определение областей с высокой концентрацией объектов одного из классов, исключение неравенств следствий и поиск эталонных точек в выбранных областях – часто используется нами при решении практических задач в различных предметных областях и подробно изложена в [5, 7]. Например, на основе данного подхода нами были сформулированы рекомендации по подбору технологических параметров для получения высококачественного агломерата на ПАО Северсталь. Анализировались 4500 наблюдений при 43 признаках. Наилучшие результаты показал КС из пяти членов. Были выделены 2 области с высокой концентрацией высококачественного агломерата. Для этих областей были найдены эталонные точки, к которым следует стремиться для получения агломерата, соответствующего высоким качественным характеристикам.

2. Проверка совместности системы неравенств и поиск максимально совместных подсистем. В предыдущем параграфе мы рассмотрели случай анализа РП, построенного для одной нозологии. На практике часто встречаются ситуации, когда найдены РП в виде линейного разделения множеств для нескольких нозологий. В нашей практике был случай, когда анализировались 124 заболевания у 6200 женщин и 62 заболевания у 1900 мужчин. Для 17 нозологий у женщин и 14 нозологий у мужчин были найдены РП с достаточными метриками качества в одном и том же пространстве признаков в виде простого линейного разделения множеств. Генеральные выборки по мужчинам и женщинам разделялись случайным образом на обучающую и тестовую в пропорции 80/20, и делалось это для каждого заболевания. Каждое такое неравенство задавало полупространство, нахождение в котором означало, что пациент здоров. Поэтому сразу возникли задачи нахождения областей, в которых отсутствовали бы все 17 заболеваний для женщин и 14 заболеваний для мужчин. Для этих целей требовалось проверить систему найденных РП на совместность, а в случае ее несовместности найти максимально совместную подсистему (МСП) неравенств. Такие задачи достаточно просто решаются, как задачи линейного программирования (ЗЛП) с частично-булевыми переменными.

Листинг 2. Модель поиска допустимого решения системы неравенств в случае ее совместности и МСП максимальной длины в случае ее несовместности

```
int n = ...; //число входных признаков
range i = 1..k; // индекс входного признака
int m = ...; //число неравенств
range j = 1..m; // индекс неравенства

float A[j][i]=...; //коэффициенты системы неравенств
float B[j]=...; //правые части системы неравенств
int L=...; // большое число (во много раз больше максимального по
абсолютной //величине B[j] и A[j][i])

/* искомые переменные */
dvar float+ x[i] ; //допустимое решение системы
dvar boolean z[j] ; //признак несовместности неравенства

minimize sum (j in j) z[j];

subject to {

forall( j in j) sum (i in i) A[j][i]*x[i]+B[j]>=-L*z[j];

};
```

В случае совместности системы неравенств $(\sum (j \text{ in } j) z[j])=0$. Если хотя бы одно $z[j]=1$, то система несовместна. В этом случае система из неравенств, у которых $z[j]=0$, будет МСП максимальной длины.

3. Проверка совместности комитетов единогласия и поиск максимально совместной подсистемы КЕ (МСПКЕ). Если РП различных нозологий представлены комитетами единогласия, то так как область допустимых решений для каждого КЕ является выпуклым множеством, а пересечение выпуклых множеств тоже является выпуклым множеством, то возникает постановка поиска максимально возможного пересечения КЕ различных нозологий. В силу громоздкости листинга программы для этого случая приведем только математическую модель, на основе которой легко создать программу в кодах IBM ILOC CPLEX преобразовав Листинг 2. Введем следующую систему обозначений:

- I – множество входных признаков;
- i – индекс входного признака $i \in I$;
- T – множество комитетов единогласия;
- t – индекс комитета единогласия $t \in T$;
- J – множество наблюдений;
- j – индекс множества наблюдений, $j \in J$;
- A_{ji}^t – коэффициенты гиперплоскостей комитета;
- B_j^t – свободные члены гиперплоскостей комитета;
- z^t – булева переменная (признак непопадания в МСП);
- L – большое число.

$$\sum_{i \in I} A_{ji}^t * x_i + B_j^t \geq L * z^t \quad j \in J^t, t \in T; \quad (4)$$

$$\min \sum_{t \in T} z^t \quad (5)$$

Данный подход использовался нами при построении РП для прогнозирования гипертензии и гипотензии по показателям биохимии и содержанию различных микроэлементов в сыворотке крови на основе анализов 1200 женщин и 700 мужчин. Цель исследования состояла в нахождении в пространстве признаков областей, в которых отсутствуют оба этих заболевания.

4. Определение экстремальности точек выпуклого множества через альтернативные системы и задачи линейного программирования. Необходимость построения выпуклых оболочек множеств достаточно часто возникает в задачах вычислительной геометрии и при решении задач классификации [5]. Существует достаточно большое количество алгоритмов построения выпуклых оболочек (ВО) [8]. К сожалению, многие алгоритмы (Грэхэма, Джарвиса, Эндрю, Чена и др.) позволяют находить ВО только для случая двухмерного пространства. Для реальных задач такие алгоритмы не подходят. В 1994 году Р. М. Pardalos, Y. Li, W.W. Hager предложили определять экстремальность точек ВО на основе задач линейного программирования [9]. Они исходили из того, что если точка множества является внутренней, то она может быть представлена, как выпуклая линейная комбинация других точек множества. Если же точка является вершиной многогранника (в их терминологии экстремальная точка), то она не может быть представлена, как ВЛК других точек, то есть ее можно представить только через саму себя. Мы предлагаем рассмотреть различные алгоритмы определения экстремальности точек ВО с точки зрения альтернативных систем.

Для формулировки понятия “альтернативные системы” воспользуемся теоремой Фаркаша [1, 2, 6]:

$$\text{либо } A * x = b \quad x \geq 0; \text{ либо } A^T * u \leq 0 \quad b^T * u > 0 \quad (6)$$

Во второй системе требуется строгое выполнение условия: $b^T * u > 0$

Теорема Фаркаша имеет простую интерпретацию: Либо вектор b можно представить, как линейную комбинацию других векторов с положительными коэффициентами, либо через нуль можно провести гиперплоскость, такую, что векторы столбцы матрицы A и вектор b будут лежать в различных полупространствах, определяемых данной гиперплоскостью.

Рассмотрим, как теорему Фаркаша можно применить для определения экстремальности точек множества. Рассмотрение начнем с одной точки. Введем следующую систему обозначений:

j – индекс точки, $j = 1, 2 \dots n$;

i – индекс координат точки, $i = 1, 2 \dots m$;

a_{ij} – i – тая координата j – ой точки ;

b_i – i – тая координата точки b проверяемой на экстремальность;

x_j – коэффициенты линейной комбинации точек .

Точка b является экстремальной относительно других точек, если следующая система несовместна:

$$\sum_{j=1}^n a_{ij} * x_j = b_i \quad i = 1, 2 \dots m \quad (7)$$

$$\sum_{j=1}^n x_j = 1 \quad (8)$$

$$x_j \geq 0 \quad j = 1, 2 \dots n \quad (9)$$

Тогда, по теореме Фаркаша, ей можно поставить в соответствие систему:

$$\sum_{i=1}^m a_{ij} * u_i + u_{m+1} \leq 0 \quad j = 1, 2 \dots n \quad (10)$$

$$\sum_{i=1}^m b_i * u_i + u_{m+1} \geq c, \quad (11)$$

где c – некоторая положительная малая константа

Отметим, что если система (10), (11) совместна при некотором $c > 0$, то, умножив правую и левую части неравенства (11) на некоторое положительное число, можно получить в правой части любое число. Поэтому неравенство (11) можно преобразовать, и в окончательном виде система альтернативная системе (7)-(9) будет выглядеть:

$$\sum_{i=1}^m a_{ij} * u_i + u_{m+1} \leq 0 \quad j = 1, 2 \dots n \quad (10)$$

$$\sum_{i=1}^m b_i * u_i + u_{m+1} \geq 1 \quad (12)$$

Системы (7)-(9) и (10), (12) обладают следующим взаимным свойством: если одна из них совместна, до другая нет.

Система (10), (12) естественным образом может быть преобразована в задачу линейного программирования (ЗЛП):

$$\sum_{i=1}^m a_{ij} * u_i + u_{m+1} \leq 0 \quad j = 1, 2 \dots n \quad (10)$$

$$\sum_{i=1}^m b_i * u_i + u_{m+1} \leq 1 \quad (13)$$

$$\max g = \sum_{i=1}^m b_i * u_i + u_{m+1} \quad (14)$$

Заметим, что условия (10), (13) совместны всегда, так как $u_i = 0 \forall i$ и $u_{m+1} = 0$ являются их решением. Условие (13) ограничили сверху и максимизируем его значение. ЗЛП (10), (13), (14) всегда имеют решение в случае, если $g = 1$, точка b является экстремальной. Если $g < 1$, то точка b является внутренней точкой ВО. Приведем листинг этой модели.

Листинг 3. Модель проверки точки множества на экстремальность

```
int m = ...; //число параметров
range i = 1..m; // индекс параметра
int n = ...; //число точек
range j = 1..n; // индекс точки
float b[i]=...; //параметры точки тестируемой на экстремальность
float A[j][i]=...; //параметры других точек, кроме тестируемой

/* искомые переменные */
```

```
dvar float u[i] ; //переменные альтернативной системы
dvar float d ; //переменная соответствующая  $u_{m+1}$ 
dvar float gf ; //значение целевой функции
```

```
maximize gf;
```

```
subject to {
```

```
forall( j in j)
```

```
sum(i in i) A[j][i]*u[i]+d <= 0;
```

```
sum(i in i) b[i]*u[i]+d<= 1;
```

```
sum(i in i) b[i]*u[i]+d == gf;
```

```
/* Если gf=1 точка b экстремальная */}
```

Конечно, условие (13) можно заменить на: $\sum_{i=1}^m b_i * u_i + u_{m+1} = 1$ (13*) и проверить систему (10), (13*) на совместность. Однако, при таком подходе выигрыша по времени счета мы не получим, но потеряем возможность построить двойственную задачу к (10), (13), (14). Если двойственные переменные обозначить, как y_j $j = 1, 2 \dots n + 1$, двойственная задача будет выглядеть следующим образом:

$$\sum_{j=1}^n a_{ij} * y_j + b_i * y_{n+1} = b_i \quad i = 1, 2 \dots m \quad (15)$$

$$\sum_{j=1}^n y_j + y_{n+1} = 1 \quad (16)$$

$$y_j \geq 0 \quad j = 1, 2 \dots n + 1 \quad (17)$$

$$\min y_{n+1} \quad (18)$$

Так как в прямой задаче нет ограничений на переменные, то в двойственной все условия записаны в виде уравнений. Точка, проверяемая на экстремальность, является экстремальной только в случае, если $y_{n+1} = 1$. Отметим, что модель (15)-(18) имеет много общего с моделью Пардалоса [9], но записана в виде, более удобном для практического использования.

Листинг 4. Проверка точки множества на экстремальность на основе модели (15)-(18)

```
int m =...; //число параметров
range i = 1..m; // индекс параметра
int n =...; //число точек
range j = 1..n; // индекс точки
float b[i]=...; //параметры точки, тестируемой на экстремальность
float A[j][i]=...; //параметры других точек, кроме тестируемой
```

```
/* искомые переменные */
```

```
dvar float+ y[j] ; //двойственные переменные
```

```
dvar float+ yy ; // целевая двойственная переменная
```

```
minimize yy;
```

```
subject to {
```

```
forall( i in i)
```

```
sum(j in j) A[j][i]*y[j]+b[i]*yy == b[i];
```

```
sum(j in j) y[j]+yy == 1;
```

```
/* если yy = 1, то точка экстремальная*/
```

```
};
```

Данный подход использовался нами при анализе влияния на различные нозологии содержания токсичных и эссенциальных микроэлементов в моче, сыворотке крови и волосах.

Заключение. Конечно, в этой статье мы привели только небольшую часть результатов из теории линейных неравенств, которые мы активно используем на практике. Целью статьи было привлечь внимание специалистов по МО к этому мощному инструментарию. Предлагаемые нами результаты позволяют избавляться от избыточных ограничений, осуществлять поиск МСП системы неравенств и МСП комитетов единогласия, осуществлять поиск экстремальных точек выпуклых оболочек различными способами.

Список источников

1. Черников С.Н. Линейные неравенства / С.Н. Черников. – М.: Наука, 1968. – 488 с.
2. Еремин И.И. Теория линейной оптимизации / И.И. Еремин. – Екатеринбург, Изд-во «Екатеринбург», 1999. – 312 с.
3. Мазуров В.Д. Комитеты системы линейных неравенств / В.Д. Мазуров, М.Ю. Хачай // Автоматика и телемеханика, 2004. – №2. – С. 43-54.
4. Мазуров В.Д. Экзистенциальные вопросы комитетных конструкций / В.Д. Мазуров // Часть II. Вестник Южно-Уральского государственного университета, 2019. – Т.19. – №1. – С. 114-120.
5. Чернавин П.Ф. Машинное обучение на основе задач математического программирования // П.Ф. Чернавин, Д.Н. Гайнанов, В.Н. Панкращенко и др. – М.: Наука, 2021. – 128 с.
6. Зоркальцев В.И. Системы линейных неравенств. Учебное пособие / В.И. Зоркальцев, М.А. Киселева. – Иркутск: ИГУ, 2007. – 99с.
7. Чернавин П.Ф. Оптимизационные модели подбора параметров технологических процессов на основе результатов машинного обучения / П.Ф. Чернавин, Н.П. Чернавин, Ф.П. Чернавин // Информационные и математические технологии в науке и управлении, 2023. – № 2(30). – С. 45-56.
8. Вычислительная геометрия. – URL: www.kpfu.ru/staf_files/a793539611/Lecture3.pdf (дата обращения 01.07.2023)
9. Pardalos P.M., Li Y., Hager W.W. Linear programming approaches to the convex hull problem in R^m. Computer V^Fthematics Application, 1995, vol. 29, no. 7. pp. 23-29

Чернавин Павел Федорович. К.э.н., Уральский федеральный университет, доцент кафедры Аналитика больших данных и методы видеоанализа. Научные интересы: построение математических моделей с использованием методов исследования операций и машинного обучения. AuthorID: 117430, SPIN:6370-8103, ORCID: 0000-0003-3214-3906, chernavin.p.f@gmail.com, 620002, Свердловская область, г. Екатеринбург, ул. Мира д. 19.

Чернавин Николай Павлович. Уральский федеральный университет, ассистент кафедры Аналитика больших данных и методы видеоанализа, AuthorID : 971565, SPIN:5722-9436, ORCID: 0000-0003-3214-3906, ch_k@mail.ru, 620002, Свердловская область, г. Екатеринбург, ул. Мира д. 19.

Чернавин Федор Павлович. К.э.н., Уральский федеральный университет, доцент кафедры Моделирование управляемых систем, AuthorID: 971595, SPIN: 9237-5190, ORCID: 0000-0003-3214-3906, chernavin_fedor@mail.ru, 620002, Свердловская область, г. Екатеринбург, ул. Мира д. 19.

UDC 004.855.5

DOI:10.25729/ESI.2023.32.4.002

Application of the theory of linear inequalities in machine learning problems

Pavel F. Chernavin, Nikolai P. Chernavin, Fedor P. Chernavin

Ural Federal University, Russia, Yekaterinburg, p.f.chernavin@urfu.ru

Abstract. According to the authors, the results of the theory of linear inequalities should be more widely used in machine learning problems. To eliminate redundant inequalities when constructing ensembles based on linear separators, one should use theorems on dependent inequalities and consequences. To generalize the results of various studies, search models for the most compatible subsystems should be used. The concept of maximally collaborative subsystems should be expanded to include unanimity committees. To solve classification problems,

one can effectively use the method of convex hulls, and to determine the extreme points of convex hulls, use the results from the theory of alternative systems. The article provides the information necessary for this from the theory of linear inequalities, mathematical models based on them and a listing of programs for the computer implementation of mathematical models.

Keywords: machine learning, linear inequalities, maximally consistent subsystems, mathematical programming, classification, medical diagnostics

References

1. Chernikov S.N. Lineynyye neravenstva [Linear inequalities]. M., Nauka [Science], 1968, 488 p.
2. Yeregin I.I. Teoriya lineynoy optimizatsii [Theory of linear optimization]. Yekaterinburg, Izd-vo "Yekaterinburg" [Yekaterinburg Publishing House], 1999, 312 p.
3. Mazurov V.D., Khachai M.Yu. Komitety sistemy lineynykh neravenstv [Committees of the system of linear inequalities]. Avtomatika i telemekhanika [Automation and telemekhanics], 2004, no.2, pp. 43-54.
4. Mazurov V.D. Ekzistencial'nye voprosy komitetnykh konstruktsiy [Existential questions of committee constructions]. Chast' II. Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta [Part II. Bulletin of the South Ural state university], 2019, v.19, no.1, pp. 114-120.
5. Chernavin P.F., Gajnanov D.N., Pankrashchenko V.N. et al. Mashinnoe obuchenie na osnove zadach matematicheskogo programmirovaniya [Machine learning based on mathematical programming problems]. Moscow, Nauka [Science], 2021, 128 p.
6. Zorkal'tsev V.I., Kiseleva M.A. Sistemy lineynykh neravenstv. Uchebnoye posobiye [Systems of linear inequalities. Tutorial]. Irkutsk ISU, 2007, 99 p.
7. Chernavin P.F. Optimizatsionnyye modeli podbora parametrov tekhnologicheskikh protsessov na osnove rezul'tatov mashinnoy obucheniya [Optimization models for selecting technological process parameters for based on machine learning results]. Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii [Information and mathematical technologies in science and management], 2023, no. 2(30), pp. 45-56.
8. Vychislitel'naya geometriya. [Computational geometry]. Available at: www.kpfu.ru/staf_files/a793539611/Lecture3.pdf (accessed: 07/01/2023)
9. Pardalos P.M., Li Y., Hager W.W. Linear programming approaches to the convex hull problem in R^m. Computer VFIematics Application, 1995, vol. 29, no. 7. pp. 23-29

Chernavin Pavel Fedorovich. Ph.D., Ural federal university, associate professor of the department of Big data analytics and video analysis methods. Research interests: building mathematical models using operations research and machine learning methods. AuthorID: 117430, SPIN: 6370-8103, ORCID: 0000-0003-3214-3906, chernavin.p.f@gmail.com, 620002, Russia, Sverdlovsk region, Ekaterinburg, st. Mira 19.

Chernavin Nikolay Pavlovich. Ural federal university, associate professor of the department of Big data analytics and video analysis methods. AuthorID: 971565, SPIN:5722-9436, ORCID: 0000-0003-3214-3906, ch_k@mail.ru, 620002, Russia, Sverdlovsk region, Ekaterinburg, st. Mira 19.

Chernavin Fedor Pavlovich. Ph.D., federal university, associate professor of the department of Big data analytics and video analysis methods. AuthorID: 971595, SPIN: 9237-5190, ORCID: 0000-0003-3214-3906, chernavin_fedor@mail.ru, 620002, Russia, Sverdlovsk region, Ekaterinburg, st. Mira 19.

Статья поступила в редакцию 10.07.2023; одобрена после рецензирования 06.10.2023; принята к публикации 01.12.2023.

The article was submitted 07/10/2023; approved after reviewing 10/06/2023; accepted for publication 12/01/2023.

Создание Больших Знаний и расширение областей применения миварных технологий логического искусственного интеллекта

Варламов Олег Олегович

МГТУ им. Н.Э. Баумана, НИИ МИВАР, Россия, Москва,

ovar@yandex.ru

Аннотация. Преимущества миварных экспертных систем логического искусственного интеллекта (ИИ) заключаются в кардинальном снижении вычислительной сложности автоматического построения алгоритмов и логического вывода с $N!$ до линейной N и расширении базовых продукций формата «Если, То» с формальной логики до реализации вычислительных процедур в едином информационно-управляющем пространстве. Это позволяет создавать системы поддержки принятия решений, моделировать реальные бизнес-процессы, планировать и перестраивать в реальном времени действия робототехнических комплексов и киберфизических систем, а также автоматически строить алгоритмы решения задач на основании миварной базы знаний. «Большие Знания» – это объединение и синтез разнородных баз знаний, что обеспечивает качественный переход и предоставляет большие возможности для создания систем ИИ. Разработкой машины логического вывода «Разуматор» занимаются программисты, а созданием Больших Знаний – аналитики, которых называют: инженеры знаний, когнитологи и т.п. Большие Знания можно постепенно наращивать и увеличивать в размерах, как количественно, так и качественно, например, для создания Активной Миварной Энциклопедии. Создание новых баз знаний для различных предметных областей привело к расширению областей применения миварных технологий и позволило решать новые для экспертных систем задачи: планирования маршрутов роботов; оптимизации распределения ресурсов; планирования действий и составление сметы проекта; динамический расчет многомерных векторов и их сравнение; подбора команд и комбинаций персонажей; информационной безопасности и многого других. Фактически, это все задачи, для решения которых человек рассуждает и использует правила «Если-То» или процедуры «Вход-Действие-Выход». Кроме того, для создания комплексных систем ИИ миварные технологии успешно совмещаются с нейросетевыми методами, например, для распознавания образов и речи. Логический искусственный интеллект создан на основе миварных технологий и теперь его надо обучать путем создания Больших Знаний, что позволит повысить производительность труда и создать автономных интеллектуальных роботов и многое другое.

Ключевые слова: мивар, миварные сети, базы знаний, онтологии, система принятия решений, миварная экспертная система, MOGAN, MIPRA, КЭСМИ, Wi!Mi, Разуматор, Большие Знания, роботы, моделирование, обучение

Цитирование: Варламов О.О. Создание Больших Знаний и расширение областей применения миварных технологий логического искусственного интеллекта / О.О. Варламов // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 30-41. – DOI: 10.25729/ESI.2023.32.4.003.

Введение. В настоящее время в области логического искусственного интеллекта (ИИ) продолжается активное развитие научной школы миварных технологий [1], по которым в РИНЦ размещено уже более 730 научных публикаций. В международных англоязычных журналах уровня Q1/Q2 также опубликованы статьи по созданию миварных экспертных систем (МЭС), например, конструктор МЭС – КЭСМИ, Wi!Mi, Разуматор [2], применение МЭС для беспилотного транспорта [3] и в качестве систем принятия решений – «мозгов для автономных роботов» [4]. Напомним, что главное преимущество миварных экспертных систем логического искусственного интеллекта (ЛИИ) состоит в кардинальном снижении вычислительной сложности логического вывода и/или автоматического построения алгоритмов с $N!$ до линейной (N) [1, 2] (относительно количества правил «Если, То») и расширение базовых продукций формата «Если, То» с обычной логики до реализации вычислительных процедур, сервисов, агентов и т.п., включая описание процессов в формате «Вход-Действие-Выход», в едином информационно-управляющем пространстве (ЕИУП). Это позволяет создавать системы поддержки принятия решений, моделировать реальные бизнес-процессы, планировать и перестраивать в реальном времени действия робототехнических комплексов (РТК) и киберфизических

систем, а также автоматически строить алгоритмы решения задач на основании миварной базы знаний [5].

«Большие Знания» – это объединение и синтез разнородных баз знаний, что обеспечивает качественный переход и предоставляет большие возможности для создания систем ИИ. Напомним, что изначально область ИИ занималась автоматизацией разумной деятельности человека. Все соответствующие западные и российские исследования направлены именно на это (миварные технологии преследуют ту же цель). Поэтому, несмотря на различные споры о том, что же является искусственным интеллектом, в данном контексте целесообразно будет привести следующее формальное определение ИИ: искусственный интеллект – область информатики, занимающаяся автоматизацией разумной деятельности человека [5]. Кроме того, напомним, что согласно российскому классификатору УДК, искусственный интеллект является частью кибернетики.

Системный анализ научной области ИИ позволил создать системную трехмерную модель ИИ (рис.1), которая объединяет все области, уровни исследований и типы создаваемых систем ИИ:

- представление знаний в разделе «экспертные системы»;
- смысловое понимание естественного языка в разделе семантической обработки текстов;
- распознавание образов, которые могут генерироваться различными датчиками и представлять из себя картинки, видео, звуки и т.п. сигналы от систем технического зрения в разделе «смысловое понимание образов»;
- управление своими «актерами» для роботов в реальном физическом мире в разделе «автоматизированные системы управления».

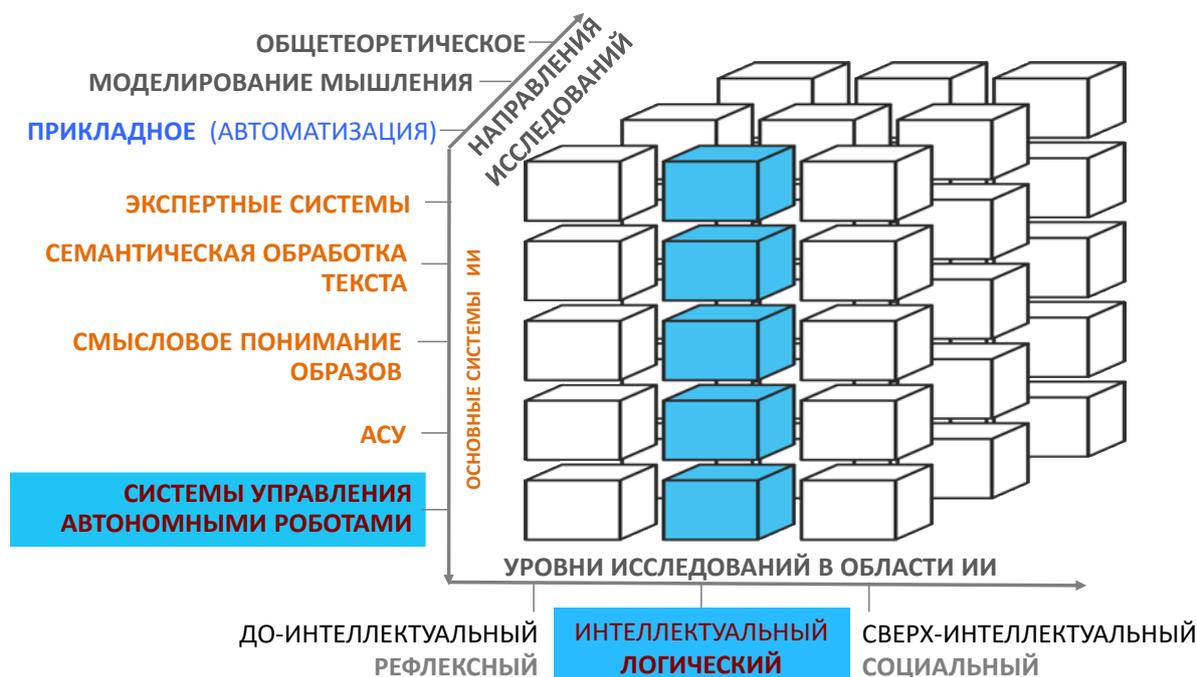


Рис. 1. Системная модель ИИ и место миварных технологий

Отдельно выделен класс систем ИИ в виде систем принятия решений и управления автономными роботами, которые должны обладать всеми предыдущими системами, да еще и работающими вместе внутри таких роботов, что исключает возможность использования суперкомпьютеров для дистанционного управления роботами.

Под термином «системы логического искусственного интеллекта» (СИИ) в этой работе понимается подкласс СИИ, которые можно охарактеризовать как активные самообучающиеся

логически рассуждающие системы (ЛРС). Стоит отметить, что представленное определение не раскрывает проблему создания полностью автоматического ИИ, который по своим возможностям будет превосходить как «типового» человека [6], так и самого «умного», и даже коллективы специалистов-экспертов, например, решающих сверхсложные задачи оценивания и прогнозирования сложных предметных областей [7].

Как известно, логический вывод и экспертные системы были основой «второй волны ИИ» в 80-90 годы прошлого века [8]. Термин «базы знаний» появился тогда же и широко использовался в области ИИ [8]. База знаний (БЗ) – основа систем логического ИИ, как одного из классов интеллектуальных систем, где знания описаны на некотором языке представления знаний [9]. Сегодня знания приобрели чисто декларативную форму, то есть знаниями считаются предложения, записанные на языках представления знаний, приближенных к естественному языку и понятных неспециалистам. Там же написано, что под здравым смыслом понимают знания, позволяющие принимать правильные решения и делать правильные предположения, основываясь на логическом мышлении и накопленном опыте [9, 10]. В наше время повсеместной «рекламы» нейросетевых технологий, целесообразно напомнить, что термин «искусственный интеллект» (artificial intelligence) был предложен в 1956 году. При этом, слово intelligence означает «умение рассуждать разумно», а вовсе не «интеллект», для которого есть термин intellect [9]. В классической книге Дж. Люгера отмечено, что интеллект – это сложная область знаний, которую невозможно описать с помощью какой-то одной теории [6]. Подчеркнем, что для создания и развития перспективных систем ИИ (примеры которых показаны в [11-23] и подробнее описаны ниже), рефлексные статистические нейросетевые технологии необходимо использовать вместе с технологиями логического уровня ИИ, к которым относятся экспертные системы.

Научная область «экспертные системы». Миварные технологии стартовали с решения задач представления знаний и их быстрой логико-вычислительной обработки, что ближе всего к области экспертных систем [1-2]. Приведем краткую информацию по тематике традиционных [24, 25] экспертных [26, 27] систем (ЭС). Экспертные системы [24, 28] как самостоятельное [29, 30] направление в искусственном интеллекте [28] сформировалось в конце 1970-х гг. Этап активного развития ЭС в СССР [31, 32] пришелся на 1980-1990-ые годы [33, 34], когда было создано [30] несколько отечественных [31, 33] ЭС [34].

Важную роль в истории ЭС сыграло сообщение японского комитета по разработке ЭВМ пятого поколения, в котором основное внимание уделялось развитию «интеллектуальных способностей» компьютеров [10], с тем, чтобы они могли оперировать не только данными, но и знаниями, как это делают специалисты (эксперты) при выработке умозаключений. Обычно к ЭС относят системы, основанные на знании, т.е. системы, вычислительная возможность которых является в первую очередь следствием их наращиваемой базы знаний (БЗ) и только во вторую очередь определяется используемыми методами. Экспертные системы не отвергают и не заменяют традиционного подхода к программированию, они отличаются от традиционных программ тем, что ориентированы на решение неформализованных задач [10]. Главное отличие ЭС от систем обработки данных состоит в том, что в них используется символичный, а не числовой способ представления данных, а в качестве методов обработки информации применяются процедуры логического вывода и эвристического поиска решений.

Там же в [10] приведено описание важного для ЭС понятия: одна из моделей представления знаний – продукционная модель. В моделях этого типа используются некоторые элементы логических и сетевых моделей. Из логических моделей заимствована идея правил вывода, которые здесь называются продукциями, а из сетевых моделей – описание знаний в виде семантической сети. В результате применения правил вывода к фрагментам сетевого описания

происходит трансформация семантической сети за счет смены ее фрагментов, наращивания сети и исключения из нее ненужных фрагментов. Таким образом, в продукционных моделях процедурная информация явно выделена и описывается иными средствами, чем декларативная информация.

В области ЭС практически сразу же начались исследования [35] по оптимизации [36] логического вывода и различного представления знаний [37], например, в виде онтологий [38]. Кроме того, были проведены работы по альтернативным [39] средствам ускорения вывода в продукционных [40] ЭС на основе Rete-сетей [41]. Отдельным направлением стали и попытки ускорить работу ЭС на основе сетей Петри [42], которые применялись для моделирования систем [43]. Отметим, что и в России [44] применяли подобные алгоритму RETE способы причинно-следственного упорядочивания [45] и организации знаний [46] для правдоподобных рассуждений [47]. Сразу отметим, что способ создания миварных сетей с линейной вычислительной сложностью логического вывода был предложен в 2002 году [1], его подробное описание в иностранных источниках размещено в 2011 году [48], а в 2015 зарегистрирован патент на изобретение автоматизированного построения маршрута логического вывода в миварной базе знаний [49]. Таким образом, миварные сети позволили в общем случае избавиться от полного перебора и придумывания различных эвристик для логического вывода на продукциях общего вида «Если, То», что позволило обрабатывать более 5 млн продукционных правил в общем случае [50] за доли секунды на обычном компьютере или ноутбуке [2, 4].

Таким образом, обобщим, что понимают под термином экспертные системы в литературе [2, 6, 8-10, 12, 16, 24-50]: экспертные системы предназначены для решения задач на основе накапливаемой базы знаний, отражающей опыт работы экспертов в рассматриваемой проблемной области. В настоящее время ЭС используются при решении задач следующих типов: принятие решений в условиях неопределенности (неполноты), интерпретация символов и сигналов, предсказание, диагностика, конструирование, планирование, управление, контроль и др. Кроме того, в [10] конкретизированы классы задач, решаемых экспертными системами реального времени (динамическими): мониторинг в реальном масштабе времени, системы управления верхнего уровня, системы обнаружения неисправностей, диагностика, составление расписаний, планирование, оптимизация, системы-советчики оператора, системы проектирования [10]. Важным ограничением ЭС в те годы было то, что алгоритмическое решение существует, но его нельзя использовать по причине большой размерности пространства решений и ограничений на ресурсы (времени, памяти) [10].

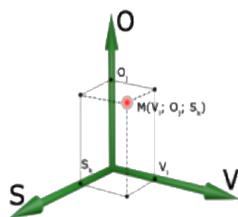
Миварный линейный логический вывод и переход «количества в качество». Напомним, что разработаны три миварные технологии [1, 5, 11] (рис.2), которые получены в результате фундаментальных научных исследований и доведены до практического использования [12] и внедрений в различных программно-аппаратных комплексах [13]:

1. Накопление информации в формализме гносеологической модели знаний «Вещь-Свойство-Отношение» (ВСО), которая обобщает все известные модели данных баз данных.

2. Логический вывод линейной сложности на миварных сетях «переменные, правила» ($\langle P, R \rangle$), созданных путем обобщения продукционного подхода и сетей Петри.

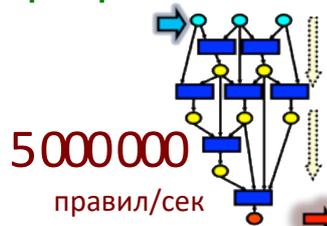
3. Многомерная открытая гносеологическая активная сеть (MOGAN), позволяющая проводить и накопление, и логико-вычислительную обработку на миварных сетях в ВСО.

1. Единое Информационное Пространство



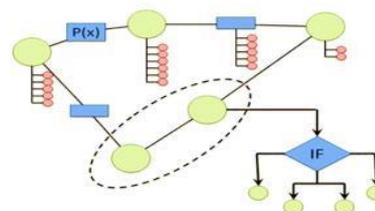
**НАКОПЛЕНИЕ
ИНФОРМАЦИИ=>
БОЛЬШИЕ ЗНАНИЯ**

2. Единое Управляющее Пространство



**БЫСТРАЯ ЛОГИЧЕСКАЯ И
ВЫЧИСЛИТЕЛЬНАЯ
ОБРАБОТКА ИНФОРМАЦИИ**

3. Единое Информационно-Управляющее Пространство



**МНОГОМЕРНАЯ ОТКРЫТАЯ
ГНОСЕОЛОГИЧЕСКАЯ
АКТИВНАЯ СЕТЬ (MOGAN)**

БАЗЫ ЗНАНИЙ «VSO», ЛИНЕЙНЫЙ ЛОГИЧЕСКИЙ ВЫВОД, СТАТИСТИЧЕСКАЯ И ВЫЧИСЛИТЕЛЬНАЯ ОБРАБОТКА ПРЕДСТАВЛЯЮТ СОБОЙ ЕДИНОЕ ЦЕЛОЕ, ЧТО ПОЗВОЛЯЕТ АВТОМАТИЧЕСКИ СТРОИТЬ **ГИБКИЕ АЛГОРИТМЫ РЕШЕНИЙ** И ПРОИЗВОДИТЬ ВЫЧИСЛЕНИЯ В АВТОНОМНЫХ РОБОТАХ

Рис. 2. Три основных миварных технологии логического ИИ

После снятия фундаментального ограничения по скорости логико-вычислительной обработки – «проклятия размерности полного перебора» и достижения в МЭС скорости обработки более 5 млн правил/сек [2, 5, 11-13, 48-50] на обычных однопроцессорных компьютерах, актуальной остается проблема создания различных баз знаний для самых разных предметных областей и объединение их на основе формализма миварных сетей - двудольных ориентированных графов $\langle P, R \rangle$ [1], в общую систему представления и обработки знаний.

Большие Знания. Данную задачу целесообразно называть «Большие Знания», например, по аналогии с «Большими Данными» для статистических нейросетевых методов ИИ. Здесь мы действительно видим явный переход «количества в качество», когда значительное количественное увеличение возможностей компьютеров привело к качественному росту возможностей нейросетей рефлексного ИИ. Новая научная проблема выявлена в процессе развития миварных технологий логического ИИ, которая заключается в необходимости создания разнообразных баз знаний экспертных систем в формализме миварных сетей (двудольных ориентированных графов $\langle P, R \rangle$).

Для большей наглядности этой научной проблемы, можно провести следующую аналогию: создание паровоза и строительства сети железных дорог, где миварная машина логического вывода «Разуматор» - это *паровоз*, а *сеть железных дорог* – это созданные базы знаний в самых разных предметных областях, которые в итоге и образуют «Большие Знания». По этой аналогии понятно, что разработка программного обеспечения «Разуматор» и создание баз знаний Больших Знаний – это две разные области деятельности, которые требуют и разных специализаций. Разработкой Разуматора занимаются программисты, а созданием Больших Знаний – аналитики, которых в данный момент называют по-разному: инженеры знаний, когнитологи и т.п. Хотя и те, и другие специальности – это все разные специализации одной «классической специальности»: прикладной математики. Миварный подход обладает свойством эволюционного развития и поэтому всегда можно добавить или изменить правила в Миварной Базе Знаний (МБЗ). По приведенной выше аналогии это означает, что можно строить различные «ж/д дороги» в разных областях научной деятельности, а потом их можно будет соединить в одну общую сеть и получить Большие Знания. Кроме того, созданные Большие Знания [14] можно постепенно наращивать и увеличивать в размерах [15], как количественно, так и качественно [16, с.236-243], например, для создания Активной Миварной Энциклопедии [16, с.178-186]. Это подтверждает вывод: логический искусственный интеллект [16] создан на основе миварных сетей и теперь его надо обучать [16, 50] путем создания Больших Знаний.

Автоматическое создание баз знаний. В последнее время возникло новое направление ИИ по автоматическому созданию баз знаний (фактов и правил) [16, 50]. Проведен системный анализ этой проблемы и учеными нашей научно-исследовательской инициативы (НИИ) МИВАР предложены новые подходы по обучению и автоматизации процессов создания Больших Знаний. Выделены два новых направления:

- создание баз знаний [16, с.194-212] на основе систем понимания естественного языка для автоматического преобразования текстов в факты и правила, т.е. в Большие Знания;
- автоматическое построение баз знаний [17] для планирования действий роботов [18] на основе генерации правил по закономерностям и возможным действиям РТК.

Отметим, что остается традиционный подход по «ручному созданию» баз знаний аналитиками-когнитологами и экспертами в различных предметных областях [16-23, 50].

Новые области применения миварных технологий ЛИИ. Кроме традиционных областей применения ЭС, учеными миварной научной школы найдены *новые предметные области*, в которых миварные технологии применяют для решения следующих научных задач:

- 1) планирования трехмерных маршрутов роботов с учетом препятствий [19];
- 2) оптимизация [16, с. 213-219] с построением нескольких логических выводов [20] на одной сети миварных правил и расчет возможных ограничений при планировании заказов;
- 3) планирование действий по обработке и сборке изделий [16, с. 420-427], а также автоматическое составление сметы для разных задач, например, строительство, выполнение проекта в ИТ-сфере [16, с. 6-12];
- 4) принятие обоснованного решения на основе гибкого расчета, получаемого путем логического вывода, многомерных векторов и их динамического сравнения [21] в реальном времени, например, при покупке рекламы [22], оценки соотношения «покупатель» и «товар», выборе «целевой группы» и т.п. [16, с. 68-74];
- 5) для автоматизации различных служб технической поддержки [16, с. 20-28];
- 6) для технической [16, с. 75-81] защиты информации [16, с. 269-275] и информационной безопасности [16, с. 29-37] различных организаций [16, с. 59-67];
- 7) подбора команд [16, с. 122-129] и комбинаций различных персонажей [16, с. 44-53] по разным критериям;
- 8) управления образовательными программами [16, с. 323-329] и учебными дисциплинами [16, с. 330-337];
- 9) решения кадровых вопросов [16, с. 98-104] и уменьшения текучести кадров [16, с. 171-177];
- 10) поиска информации по репозиториям [16, с. 165-171] и библиотекам [16, с. 236-243], а также для многого другого, изложенного в [16, 50].

Важно подчеркнуть, что в настоящее время неоднократно показано [16, 50] совместное [16, с. 244-250] применение миварных технологий логического ИИ с нейросетевыми методами распознавания образов при работе с трехмерными объектами [23], распознавания речи [16, с. 275-280] и дорожных знаков [16, с. 302-308].

Таким образом, количественный рост скорости логического вывода и расширение формализма представления Больших Знаний позволяют перейти к новому качеству и значительно расширить области применения миварных технологий логического ИИ [11-23].

Заключение. Достоинства миварных экспертных систем логического искусственного интеллекта заключаются в кардинальном снижении вычислительной сложности логического вывода с $N!$ до линейной (N) и расширении базовых продукций формата «Если, То» с обычной логики до реализации вычислительных процедур в едином информационно-управляющем

пространстве. Разработкой Разуматора занимаются программисты, а созданием Больших Знаний – аналитики, которых называют: инженеры знаний, когнитологи и т.п. Большие Знания можно постепенно наращивать и увеличивать в размерах, как количественно, так и качественно, например, для создания Активной Миварной Энциклопедии.

Расширение областей применения миварных технологий позволило решать новые для экспертных систем задачи: планирование действий и составление сметы проекта; динамический расчет многомерных векторов и их сравнение; подбор команд и комбинаций персонажей; информационная безопасность и многое другое. Фактически, это все задачи, для решения которых человек задумывается и рассуждает.

Для создания комплексных систем ИИ миварные технологии успешно совмещаются с нейросетевыми методами, например, для распознавания образов и речи. Логический ИИ создан на основе миварных сетей и теперь его надо обучать путем создания Больших Знаний.

Термин «Большие Знания» означает не только рост количества информации, но качественное возрастание возможностей систем ИИ с использованием миварных технологий.

Список источников

1. Варламов О.О. Эволюционные базы данных и знаний для адаптивного синтеза интеллектуальных систем. / О.О. Варламов // Миварное информационное пространство. – М.: Радио и связь, 2002. – 288 с. – EDN: RWTCOP.
2. Varlamov O.O. Wi!Mi expert system shell as the novel tool for building knowledge-based systems with linear computational complexity. International review of automatic control, 2018, vol. 11, no. 6, pp. 314-325, EDN: NUCCC.
3. Shadrin S.S., Ivanov A.M., Varlamov O.O. Experimental autonomous road vehicle with logical artificial intelligence. Journal of advanced transportation, 2017, vol. 2017, pp. 2492765, EDN: XNTRKD.
4. Varlamov O. “Brains” For Robots: Application of the mivar expert systems for implementation of autonomous intelligent robots. Big Data Research, 2021, vol. 25, pp. 100241, EDN: HRYLWL.
5. Варламов О.О. Основы создания миварных экспертных систем. Москва, 2021, EDN: GQYNWE.
6. Люгер Дж.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. 4-е изд. / Дж.Ф. Люгер. – М.: Вильямс, 2005. – 864 с.
7. Ростовцев Ю.Г. Математические методы и модели оценивания военно-политической обстановки / Ю.Г. Ростовцев. – МО СССР, 1986. – 312 с.
8. Гаврилова Т.А. Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский. – Санкт-Петербург: Питер, 2000. – 384 с. – (Учебник). – ISBN 5-272-00071-4. – EDN: TGWANT.
9. Остроух А.В. Введение в искусственный интеллект / А.В. Остроух. – Красноярск: Научно-инновац. центр, 2020. – 250 с.
10. Остроух А.В. Интеллектуальные системы / А.В. Остроух. – Красноярск: Научно-инновационный центр, 2020. – 316 с.
11. Варламов О.О. Миварные базы данных и правил / О.О. Варламов. – Москва, 2021. – EDN: CGGSAQ.
12. Варламов О.О. 18 примеров миварных экспертных систем / О.О. Варламов. – Москва, 2021. – EDN: HAMRDZ.
13. Варламов О.О. Обзор 18 миварных экспертных систем, созданных на основе MOGAN / О.О. Варламов // Известия Кабардино-Балкарского научного центра РАН, 2021. – № 3 (101). – С. 5-20. – EDN: PBDXHI.
14. Бадалов А.Ю. Активная миварная интернет-энциклопедия и развитие миварных сетей на основе многомерных бинарных матриц / А.Ю. Бадалов, Р.А. Санду // Искусственный интеллект, 2010. – № 4. – С. 549-557. – EDN: UIMXAV.
15. Варламов О.О. Миварный подход к разработке интеллектуальных систем и проект создания мультипредметной активной миварной интернет-энциклопедии / О.О. Варламов. // Известия Кабардино-Балкарского научного центра РАН, 2011. – № 1-1 (39). – С. 55-64. – EDN: NECGIB.
16. Мивар'22. – Москва: Издательский Дом «Инфра-М», 2022. – 440 с. – EDN: RQIFBK.
17. Коценко А.А. О подходе к автоматической генерации миварных баз знаний для обеспечения трехмерного движения роботов / А.А. Коценко // В сборнике: Перспективные направления развития отделочно-упрочняющих и виброволновых технологий, 2023. – С. 301-308. – EDN: RKVDHK.
18. Аладин Д.В. О подходе автоматического построения миварных баз знаний для решения управленческих задач в технических системах / Д.В. Аладин // В сборнике: Перспективные направления развития отделочно-упрочняющих и виброволновых технологий, 2023. – С. 208-216. – EDN: JTOVHK.

19. Коценко А.А. О подходе к применению миварных экспертных систем для планирования трехмерных маршрутов роботов с учетом препятствий / А.А. Коценко // В сборнике: Перспективные направления развития отделочно-упрочняющих и виброволновых технологий, 2023. – С.309-315. – EDN: GFBRKU.
20. Коценко А.А. Методика применения миварной экспертной системы для автоматизированного поиска нескольких траекторий робота / А.А. Коценко, А.В. Герасименко, А.В. Калашникова и др. // Естественные и технические науки, 2022. – № 5(168). – С. 209-221. – DOI 10.25633/ETN.2022.05.11. – EDN: XSXYXM.
21. Семенов А.А. Разработка метода сравнения двух многомерных векторов в реальном времени на основе миварных экспертных систем / А.А. Семенов // Проблемы искусственного интеллекта, 2020. – № 2 (17). – С. 94-109. – EDN: FXAOKP.
22. Семенов А.А. Исследование способов подбора рекламных кампаний на основе сравнения многомерных векторов / А.А. Семенов // Проблемы искусственного интеллекта, 2020. – № 1 (16). – С. 89-104. – EDN: UEBEPL.
23. Синцов М.Ю. О развитии миварного подхода к интеллектуальному распознаванию образов для работы с трехмерными объектами / М.Ю. Синцов, А.Ю. Озерин, А.А. Кузин [и др.]. // Радиопромышленность, 2015. – № 3. – С. 172-183. – EDN: UQEP0Z.
24. Уотермен Д. Руководство по экспертным системам / Д. Уотермен. – М.: Мир, 1989. – 388 с.
25. Таунсенд К. Проектирование и программная реализация экспертных систем на персональных ЭВМ / К. Таунсенд, Д. Фохт. – М.: Финансы и статистика, 1990. – 320 с.
26. Джарратано Д., Райли Г. Экспертные системы: принципы разработки и программирование. 4-е изд.: Пер с англ. – Москва: Вильямс, 2007. – 1152 с. – EDN: QMQTDD.
27. Giarratano J.C., Riley G.D. Expert systems: principles and programming, 4-th edition. Course technology, 2004.
28. Luger F.G. Artificial Intelligence: structure and strategies for complex problem solving. Boston, Pearson, 2009.
29. Попов Э.В. Искусственный интеллект: экспертные системы / Э.В. Попов. – М.: Наука, 1990.
30. Экспертные системы. Базы знаний и данных. Материалы семинара // Под рук. Э.В. Попова. – М.: ЦРДЖ, 1992. – 171 с.
31. Кирсанов Б.С. Отечественные оболочки экспертных систем / Б.С. Кирсанов, Э.В. Попов // Справочник по искусственному интеллекту. – М.: Радио и связь, 1990. – Т. 1. – С. 369-388.
32. Попов Э.В. Статистические и динамические экспертные системы / Э.В. Попов, И.Б. Фоминых, Е.Б. Кисель и др. – М.: Финансы и статистика, 1996.
33. Яшин А.М. Разработка экспертных систем / А.М. Яшин. – Л.: ЛПИ, 1990.
34. Справочник по искусственному интеллекту в 3-х т.//Под ред. Э.В. Попова, Д.А. Поспелова. – М.: Радио и связь, 1990.
35. Тейз А. Логический подход к искусственному интеллекту: от классической логики к логическому программированию/ А. Тейз, П. Грибомон, Ж. Луи [и др.]. – М.: Мир, 1990. – 432 с.
36. Chandru V., Hooker J. Optimization Methods for Logical Inference, New York: Wiley, 1999.
37. Russel S.J., Norvig P. Artificial Intelligence: a modern approach (3rd ed.). Boston, Pearson, 2010.
38. Sadik A.R., Urban B. An ontology-based approach to enable knowledge representation and reasoning. Future Internet, 9(4), 90, 2017.
39. Forgy Ch. A network match routine for production systems. Working Paper, 1974.
40. Forgy Ch. On the efficient implementation of production systems. Ph. D. Thesis. Carnegie-Mellon University, 1979.
41. Forgy Ch. Rete: A fast algorithm for the many pattern/many object pattern match problem. Artificial intelligence, 1982, 19, pp. 17-37.
42. Котов В.Е. Сети Петри / В.Е. Котов. – М: Наука, 1984. – 160 с.
43. Питерсон Дж. Теория сетей Петри и моделирование систем / Дж. Питерсон. – М.:, 1984. – 264с.
44. Аршинский Л.В. Многозначные логики с векторной семантикой / Л.В. Аршинский. – Восточно-Сибирский институт МВД России. – Иркутск, 2003. – 46 с. – Деп. в ВИНТИ 13.02.2003, № 281-B2003. – EDN: QJBXBJ.
45. Аршинский Л.В. Интервальное оценивание истинности в системах автоматизированных рассуждений на основе VTF-логик / Л.В. Аршинский // IV международная конференция «Идентификация систем и задачи управления» SICPRO'05: труды. Москва. ИПУ РАН им. В.А. Трапезникова, 2005. – С. 1061-1074.
46. Аршинский Л.В. Организация знаний в системе моделирования правдоподобных рассуждений «Гераклит» / Л.В. Аршинский // XVII Байкальская Всероссийская конференция «Информационные и математические технологии в науке и управлении»: труды. Т. 3. Иркутск, ИСЭМ СО РАН, 2012. – С. 19 - 25.
47. Аршинский Л.В. Особенности работы машины вывода системы моделирования правдоподобных рассуждений "Гераклит" // Л.В. Аршинский // Информационные и математические технологии в науке и управлении, 2016. – № 2. – С. 18-29. – EDN WZBLJV.

48. Varlamov O. MIVAR: Transition from productions to bipartite graphs MIVAR nets and practical realization of automated constructor of algorithms handling more than three million production rules. Arxiv.org: arXiv preprint arXiv:1111.1321, 2011, 23 p., available at: <https://arxiv.org/abs/1111.1321>.
49. Варламов О.О. Автоматизированное построение маршрута логического вывода в миварной базе знаний / О.О. Варламов, А.М. Хадиев, М.О. Чибирова [и др.] // Патент на изобретение RUS 2607995 11.02.2015., опубликовано 11.01.2017. – бюл. №2. – 43 с.
50. Мивар'23: Сборник статей. – М. "Инфра-М", 2023. – 598 с. – EDN: NXALJA.

Варламов Олег Олегович. Профессор, доктор технических наук, МГТУ им. Н.Э. Баумана, профессор кафедры ИУ5 Системы обработки информации и управления, научный руководитель Научно-исследовательской инициативы (НИИ) МИВАР на кафедре ИУ5 в МГТУ им. Н.Э. Баумана. Основные направления исследований автора: комплексные системы искусственного интеллекта, кибернетика, информатика, миварные технологии логического искусственного интеллекта, базы знаний, экспертные системы, системы поддержки принятия решений, понимание текстов на русском языке, понимание образов, системы принятия решений для автономных робототехнических комплексов. AuthorID: 143177; SPIN: 7983-9762; Google Scholar: kDy5G1MAAAJ; ORCID: 0000-0002-2858-1383; Scopus: 57195354995; ResearcherID: U-4856-2017, ovar@yandex.ru, ovarlamov@gmail.com, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1.

UDC 004.891 + 004.896 + 007.52

DOI: 10.25729/ESI.2023.32.4.003

Creating Big Knowledge and expanding the applications of mivar technologies of logical artificial intelligence

Oleg O. Varlamov

BMSTU, SI MIVAR, Russia, Moscow, ovarlamov@gmail.com

Abstract. The advantages of mivar expert systems of logical artificial intelligence (AI) lie in the drastic reduction in the computational complexity of automatically building algorithms and inference from $N!$ to linear N and expansion of basic productions of the “If, Then” format from formal logic to the implementation of computational procedures in a single information and control space. This allows you to create decision support systems, simulate real business processes, plan and rebuild the actions of robotic complexes and cyber-physical systems in real time, and automatically build algorithms for solving problems based on the mivar knowledge base. “Big Knowledge” is the combination and synthesis of heterogeneous knowledge bases, which provides a qualitative transition and provides great opportunities for creating AI systems. The development of the inference machine “Razumator” is done by programmers, and the creation of Big Knowledge is done by analysts, who are called: knowledge engineers, cognitologists, etc. Great Knowledge can be gradually increased and increased in size, both quantitatively and qualitatively, for example, to create an Active Mivar Encyclopedia. The creation of new knowledge bases for various subject areas led to the expansion of the areas of application of mivar technologies and made it possible to solve new problems for expert systems: planning robot routes; optimization of resource allocation; action planning and project budgeting; dynamic calculation of multidimensional vectors and their comparison; selection of teams and combinations of characters; information security and many others. In fact, these are all problems for which a person reasons and uses the If-Then rules or the Entry-Action-Exit procedures. In addition, to create complex AI systems, mivar technologies are successfully combined with neural network methods, for example, for image and speech recognition. Logical artificial intelligence was created on the basis of mivar technologies and now it must be trained by creating Big Knowledge, which will increase labor productivity and create autonomous intelligent robots and much more.

Keywords: mivar, mivar networks, knowledge bases, ontologies, decision-making system, mivar expert system, MOGAN, MIPRA, CESMI, Wi!Mi, Razumator, Big Knowledge, robots, modeling, learning

References

1. Varlamov O.O. Evolyucionnye bazy dannyh i znaniy dlya adaptivnogo sinteza intellektual'nyh sistem. Mivarnoe informacionnoe prostranstvo [Evolutionary databases and knowledge for adaptive synthesis of intelligent systems. Mivar information space]. M., Radio i svyaz' [Radio and communication], 2002, 288 p., EDN RWTCOP.
2. Varlamov O.O. Wi!Mi expert system shell as the novel tool for building knowledge-based systems with linear computational complexity. International review of automatic control, 2018, vol. 11, no. 6, pp. 314-325, EDN: NUCCC.
3. Shadrin S.S., Ivanov A.M., Varlamov O.O. Experimental autonomous road vehicle with logical artificial intelligence. Journal of advanced transportation, 2017, vol. 2017, pp. 2492765, EDN: XNTRKD.
4. Varlamov O. "Brains" For Robots: Application of the mivar expert systems for implementation of autonomous intelligent robots. Big Data Research, 2021, vol. 25, pp. 100241, EDN: HRYLWL.
5. Varlamov O.O. Osnovy sozdaniya mivarnyh ekspertnyh sistem [Fundamentals of creating mivar expert systems]. Moscow, 2021, EDN: GQYNWE.
6. Luger J.F. Iskusstvennyj intellekt: strategii i metody resheniya slozhnyh problem [Artificial Intelligence: strategies and methods for solving complex problems]. 4th ed. M.: Williams, 2005, 864 p.
7. Rostovtsev Yu.G. Matematicheskie metody i modeli ocenivaniya voenno-politicheskoy obstanovki [Mathematical methods and models for assessing the military-political situation]. USSR Ministry of Defense, 1986, 312 p.
8. Gavrilova T.A., Khoroshevsky V.F. Bazy znaniy intellektual'nyh sistem [Knowledge bases of intelligent systems]. St. Petersburg, Peter, 2000, 384 p., ISBN 5-272-00071-4, EDN TGWANT.
9. Ostroukh A.V. Vvedenie v iskusstvennyj intellekt [Introduction to artificial intelligence]. Krasnoyarsk, Scientific-innovative center, 2020, 250 p.
10. Ostroukh A.V. Intellektual'nye sistemy [Intelligent systems]. Krasnoyarsk Nauchno-innovats. tsentr, [Scientific and innovation center], 2020, 316 p.
11. Varlamov O.O. Mivarnye bazy dannyh i pravil [Mivar databases and rules]. Moscow, 2021, EDN: CGGSAQ.
12. Varlamov O.O. 18 primerov mivarnyh ekspertnyh sistem [18 examples of mivar expert systems]. Moscow, 2021, EDN: HAMRDZ.
13. Varlamov O.O. Obzor 18 mivarnyh ekspertnyh sistem, sozdannyh na osnove MOGAN [Review of 18 mivar expert systems created on the basis of MOGAN]. Izvestiya Kabardino-Balkarskogo nauchnogo tsentra [News of the Kabardino-Balkarian scientific center of the RAS], 2021, no. 3 (101), pp. 5-20, EDN: PBDXHI.
14. Badalov A.Yu., Sandu R.A. Aktivnaya mivarnaya internet-enciklopediya i razvitie mivarnyh setej na osnove mnogomernykh binarnykh matric [Active mivar Internet encyclopedia and development of mivar networks based on multidimensional binary matrices]. Iskusstvennyj intellekt [Artificial intelligence], 2010, no. 4, pp. 549-557, EDN: UIMXAV.
15. Varlamov O.O. Mivarnyj podhod k razrabotke intellektual'nyh sistem i proekt sozdaniya mul'tipredmetnoj aktivnoj mivarnoj internet-enciklopedii [Mivar approach to the development of intelligent systems and the project of creating a multi-subject active mivar Internet encyclopedia]. Izvestiya Kabardino-Balkarskogo nauchnogo tsentra [News of the Kabardino-Balkarian scientific center of the RAS], 2011, no. 1-1 (39), pp. 55-64, EDN: NECGIB.
16. Mivar'22 [Mivar'22]. Moskva, Izdatel'skiy Dom «Infra-M» [Publishing House "Infra-M"], 2022, 440 p., EDN RQIFBK.
17. Kotsenko A.A. O podhode k avtomaticheskoy generacii mivarnyh baz znaniy dlya obespecheniya trekhmernogo dvizheniya robotov [On an approach to the automatic generation of mivar knowledge bases to ensure three-dimensional movement of robots]. V sbornike: Perspektivnyye napravleniya razvitiya otdelochno-uprochnyayushchikh i vibrovolnovykh tekhnologiy [In the collection: Prospective directions for the development of finishing-hardening and vibration-wave technologies], 2023, pp. 301-308, EDN: RKVDHK.
18. Aladin D.V. O podhode k avtomaticheskogo postroeniya mivarnyh baz znaniy dlya resheniya upravlencheskih zadach v tekhnicheskikh sistemah [On the approach to automatically constructing mivar knowledge bases for solving management problems in technical systems]. V sbornike: Perspektivnyye napravleniya razvitiya otdelochno-uprochnyayushchikh i vibrovolnovykh tekhnologiy [In the collection: Prospective directions for the development of finishing-hardening and vibration-wave technologies], 2023, pp. 208-216, EDN: JTOVHK.
19. Kotsenko A.A. O podhode k primeniyu mivarnyh ekspertnyh sistem dlya planirovaniya trekhmernykh marshrutov robotov s uchetom prepyatstvij [On the approach to the use of mivar expert systems for planning three-dimensional routes of robots taking into account obstacles]. V sbornike: Perspektivnyye napravleniya razvitiya otdelochno-uprochnyayushchikh i vibrovolnovykh tekhnologiy [In the collection: Promising directions for the development of finishing-hardening and vibration-wave technologies], 2023, pp. 309-315, EDN: GFBRKU.
20. Kotsenko A. A., Gerasimenko A. V., Kalashnikova A. V. et al. Metodika primeneniya mivarnoj ekspertnoj sistemy dlya avtomatizirovannogo poiska neskol'kih traektorij robota [Methodology for using a mivar expert system for

- automated search of several robot trajectories]. *Yestestvennyye i tekhnicheskiye nauki* [Natural and technical sciences], 2022, no. 5(168), pp. 209-221, DOI 10.25633/ETN.2022.05.11, EDN: XSXYXM.
21. Semenov A.A. Razrabotka metoda sravneniya dvuh mnogomernykh vektorov v real'nom vremeni na osnove mivarnykh ekspertnykh sistem [Development of a method for comparing two multidimensional vectors in real time based on mivar expert systems]. *Problemy iskusstvennogo intellekta* [Problems of artificial intelligence], 2020, no. 2 (17), pp. 94-109, EDN: FXAOKP.
 22. Semenov A.A. Issledovanie sposobov podbora reklamnykh kampanij na osnove sravneniya mnogomernykh vektorov [Research on methods for selecting advertising campaigns based on comparison of multidimensional vectors]. *Problemy iskusstvennogo intellekta* [Problems of artificial intelligence], 2020, no. 1 (16), pp. 89-104, EDN: UEBEPL.
 23. Sintsov M.Yu., Ozerin A.Yu., Kuzin A.A. and others. O razvitiy mivarnogo podhoda k intellektual'nomu raspoznavaniyu obrazov dlya raboty s trekhmernymi ob"ektami [On the development of the mivar approach to intelligent pattern recognition for working with three-dimensional objects]. *Radiopromyshlennost'* [Radio industry], 2015, no. 3, pp. 172-183, EDN: UQEPOZ.
 24. Waterman D. *Rukovodstvo po ekspertnym sistemam* [Guide to expert systems]. M., Mir, 1989, 388 p.
 25. Townsend K., Focht D. *Proektirovanie i programmnaya realizaciya ekspertnykh sistem na personal'nyh EVM* [Design and software implementation of expert systems on personal computers]. M., *Finansy i statistika* [Finance and Statistics], 1990, 320 p.
 26. Giarratano D., Riley G. *Ekspertnye sistemy: principy razrabotki i programmirovaniya* [Expert systems: principles of development and programming]. Moscow, Williams, 2007, 1152 pp., EDN: QMQTDD.
 27. Giarratano J.C., Riley G.D. *Expert systems: principles and programming*, 4-th edition. Course Technology, 2004.
 28. Luger F. G. *Artificial Intelligence: Structure and Strategies for complex problem solving*. Boston, Pearson, 2009.
 29. Popov E.V. *Iskusstvennyj intellekt: ekspertnye sistemy* [Artificial intelligence: expert systems]. M., Nauka [Science], 1990.
 30. *Ekspertnye sistemy. Bazy znaniy i dannyh. Materialy seminarov* [Expert systems. Knowledge and data bases]. Seminar materials. M., TsRDZh, 1992, 171 p.
 31. Kirsanov B.S., Popov E.V. *Otechestvennyye obolochki ekspertnykh sistem. Spravochnik po iskusstvennomu intellektu* [Domestic shells of expert systems. Handbook of artificial intelligence]. M., *Radio i svyaz'* [Radio and communication], 1990, vol. 1, pp. 369-388.
 32. Popov E.V., Fominykh I.B., Kisel E.B., Shapot M.D. *Statisticheskie i dinamicheskie ekspertnye sistemy* [Statistical and dynamic expert systems]. M., *Finansy i statistika* [Finance and Statistics], 1996.
 33. Yashin A.M. *Razrabotka ekspertnykh sistem* [Development of expert systems]. L., LPI, 1990.
 34. Popov E.V., Pospelov D.A. *Spravochnik po iskusstvennomu intellektu v 3-h t.* [Handbook of artificial intelligence in 3 volumes]. M., *Radio i svyaz'* [Radio and communication], 1990.
 35. Theize A., Gribomon P., Louis J. [et al.] *Logicheskij podhod k iskusstvennomu intellektu: ot klassicheskoy logiki k logicheskomu programmirovaniyu* [Logical approach to artificial intelligence: from classical logic to logic programming]. M., Mir, 1990, 432 p.
 36. Chandru V., Hooker J. *Optimization Methods for Logical Inference*, New York, Wiley, 1999.
 37. Russel S. J., Norvig P. *Artificial Intelligence: a Modern Approach* (3rd ed.). Boston, Pearson, 2010.
 38. Sadik A. R., Urban B. An ontology-based approach to enable knowledge representation and reasoning. *Future Internet*, 9(4), 90, 2017.
 39. Forgy Ch. A network match routine for production systems. Working Paper, 1974.
 40. Forgy Ch. On the efficient implementation of production systems. Ph. D. Thesis. Carnegie-Mellon University, 1979.
 41. Forgy Ch. Rete: A fast algorithm for the many pattern/many object pattern match problem. *Artificial Intelligence*, 1982, 19, pp. 17-37.
 42. Kotov V. E. *Seti Petri* [Petri nets]. M., Nauka [Science], 1984. 160 p.
 43. Peterson J. *Teoriya setej Petri i modelirovanie sistem* [Theory of Petri nets and modeling of systems]. M., 1984, 264 p.
 44. Arshinsky L.V. *Mnogoznachnyye logiki s vektornoj semantikoj* [Multi-valued logics with vector semantics]. East Siberian Institute of the Ministry of Internal Affairs of Russia, Irkutsk, 2003, 46 p., Deposit in VINITI 02/13/2003, no. 281-B2003, EDN: QJBXBJ.
 45. Arshinsky L.V. *Interval'noe ocenivanie istinnosti v sistemah avtomatizirovannykh rassuzhdenij na osnove VTF-logik* [Interval truth estimation in automated reasoning systems based on VTF logics]. IV mezhdunarodnaya konferentsiya «Identifikatsiya sistem i zadachi upravleniya» SICPRO'05: trudy. Moskva. IPU RAN im. V.A. Trapeznikova [IV International Conference "System Identification and Control Problems" SICPRO'05: proceedings. Moscow. IPU RAS named after. V.A. Trapeznikova], 2005, pp. 1061-1074.

46. Arshinsky L.V. Organizaciya znaniy v sisteme modelirovaniya pravdopodobnyh rassuzhdenij «Geraklit» [Organization of knowledge in the system for modeling plausible reasoning “Heraclitus”]. XVII Baykal'skaya Vserossiyskaya konferentsiya «Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii»: trudy. T. 3. Irkutsk. ISEM SO RAN [XVII Baikal All-Russian Conference “Information and mathematical technologies in science and management”: proceedings. T. 3. Irkutsk. ISEM SB RAS], 2012, pp. 19-25.
47. Arshinsky L.V. Osobennosti raboty mashiny vyvoda sistemy modelirovaniya pravdopodobnyh rassuzhdenij "Geraklit" [Features of the operation of the inference engine of the system for modeling plausible reasoning "Heraclitus"]. Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii [Information and mathematical technologies in science and management], 2016, no. 2, pp. 18-29, EDN WZBLJV.
48. Varlamov O. MIVAR: Transition from productions to bipartite graphs MIVAR nets and practical realization of automated constructor of algorithms handling more than three million production rules. Arxiv.org: arXiv preprint arXiv:1111.1321, available at: <https://arxiv.org/abs/1111.1321>., 23 p., 2011.
49. Varlamov O.O. Khadiev A.M., Chibirova M.O., Sergushin G.S., Antonov P.D. Avtomatizirovannoe postroenie marshruta logicheskogo vyvoda v mivarnoj baze znaniy [Automated construction of a logical inference route in a mivar knowledge base]. Patent for invention RUS 2607995 02/11/2015, published 01/11/2017, bulletin, no. 2, 43 p.
50. Mivar'23: Sbornik statey [Mivar'23: Collection of articles]. M. "Infra-M", 2023, 598 p., EDN: NXALJA

Varlamov Oleg Olegovich. Professor, Doctor of technical sciences, BMSTU, professor of the department of Information processing and management systems (IU5). Scientific director of the Research initiative (SRI) MIVAR at the department of IU5 at BMSTU. The author's main areas of research: complex artificial intelligence systems, cybernetics, computer science, mivar technologies of logical artificial intelligence, knowledge bases, expert systems, decision support systems, understanding texts in Russian, understanding images, decision-making systems for autonomous robotic systems. AuthorID: 143177, SPIN: 7983-9762, Google Scholar: kDy5GlMAAAAJ, ORCID: 0000-0002-2858-1383, Scopus: 57195354995, ResearcherID: U-4856-2017, ovar@yandex.ru; ovarlamov@gmail.com, 105005, Moscow, 2nd Bauman'skaya st., 5, building 1.

Статья поступила в редакцию 31.07.2023; одобрена после рецензирования 04.08.2023; принята к публикации 15.12.2023.

The article was submitted 07/31/2023; approved after reviewing 08/04/2023; accepted for publication 12/15/2023.

УДК 519.854.2

DOI:10.25729/ESI.2023.32.4.004

Квазиоптимальное решение задачи коммивояжера методом эволюционного согласования

Протасов Владислав Иванович, Мирахмедов Роман Октамович, Потапова Зинаида Евгеньевна, Чернова Мария Владиславовна

Московский авиационный институт (национальный исследовательский университет),
Россия, Москва, *protonus@yandex.ru*

Аннотация. Предложен и исследован эффективный алгоритм квазиоптимального решения задачи коммивояжера коллективом акторов методом эволюционного согласования. Метод основан на использовании генетических алгоритмов. Хромосомы особей состоят из треугольников триангуляции Делоне, полученной из диаграммы Вороного. Проведено сравнение результатов работы программы, разработанной, исходя из предложенного алгоритма, с существующими известными способами.

Ключевые слова: задача коммивояжера, диаграмма Вороного, генетические алгоритмы, триангуляция Делоне, планарный граф, метод эволюционного согласования, вычислительные акторы

Цитирование: Протасов В.И. Квазиоптимальное решение задачи коммивояжера методом эволюционного согласования / В.И. Протасов, Р.О. Мирахмедов, З.Е. Потапова, М.В. Чернова // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 42-50. – DOI:10.25729/ESI.2023.32.4.004.

Введение. В данной работе рассматривается решение задачи коммивояжера (ЗК) с использованием концепции вычислительных акторов [1], метода эволюционного согласования решений (МЭС) [2] и триангуляции Делоне [3], полученной для N городов из диаграммы Вороного [4].

Задача коммивояжера является NP -полной задачей [5]. Время работы алгоритма существенно зависит от размера планарного графа, описывающего эту задачу и выбранного метода решения.

Необходимо найти эффективный алгоритм квазиоптимального решения для уменьшения количества вычислений и соответствующего увеличения числа вершин графа по сравнению с существующими методами. Задача коммивояжера является актуальной и применяется в различных сферах: при маршрутизации транспортных потоков, при выборе оптимальной траектории движения рабочего инструмента, в астрономии для уменьшения времени перемещения телескопа между объектами наведения, в конструировании микросхем и многих других приложениях [6].

Данная работа изучает возможность создания квазиоптимального метода решения задачи коммивояжера на основе концепции вычислительных акторов, МЭС и разбиения планарного графа городов с помощью диаграмм Вороного и триангуляции Делоне.

Задача коммивояжера – это классическая задача оптимизации, направленная на поиск кратчайшего возможного маршрута, которым может воспользоваться коммивояжер, чтобы посетить ряд городов и вернуться в исходную точку. Проблема была впервые представлена в 1800-х годах ирландским математиком Гамильтоном. Он был заинтересован в поиске способа решения проблемы сокращения времени на доставку почты почтальоном. Возможный путь обхода городов назван гамильтоновым циклом. Число таких циклов зависит от числа городов N и равно $N!$ Поэтому нахождение гамильтонова цикла минимальной длины является для большого количества городов задачей чрезвычайной трудности. Поскольку ЗК имеет важные и многочисленные практические приложения, то поиску методов ее решения посвящены многие исследования. Для практики достаточно находить квазиоптимальные решения – когда относительная длина полученного гамильтонова цикла не превышает 1.01-1.05 от длины минимального цикла.

В середине 20-го века ЗК привлекла широкое внимание, как теоретическая проблема в математике и информатике. В 1956 году математик Меррилл Флуд [7] поставил проблему, как математическую задачу о существовании эффективного алгоритма решения для большого числа городов.

В последующие десятилетия исследователи разработали различные подходы к решению ЗК, включая алгоритм ветвей и границ [8], алгоритм ближайшего соседа [3] и генетический алгоритм [9, 10, 11]. Однако, несмотря на большие усилия, не было найдено алгоритма с полиномиальным временем, который мог бы решить проблему точно для всех случаев.

Метод ветвей и границ был впервые предложен для решения задачи коммивояжера математиком Альбертом Уильямом Такером в 1960 году. [8] Такер также ввел термин "ветви и границы" для алгоритма решения ЗК. Метод включает в себя разбиение ЗК на более мелкие подзадачи и систематическое исследование пространства решений при отслеживании нижних границ оптимального решения. Это позволяет алгоритму обрезать ветви дерева поиска, которые не могут привести к оптимальному решению, что делает поиск более эффективным. С тех пор метод ветвей и границ стал одним из наиболее широко используемых алгоритмов для решения ЗК, а также многих других задач комбинаторной оптимизации. Несмотря на предельную простоту формулирования задачи, ЗК имеет большую вычислительную сложность, но, несмотря на это, ЗК продолжает оставаться активной областью исследований, разрабатываются новые подходы и методы для эффективного ее решения.

Одним из таких прогрессирующих методов оказался подход С.К. Лау и Л.Е. Шу [3], использующий представления вычислительной геометрии в виде триангуляции Делоне [12, 13].

Практические задачи, решаемые вышеприведенными методами, объединяют следующие свойства. Эти задачи: 1) легко формализуемы; 2) представимы графом; 3) имеют только детерминированные переменные; 4) имеют один тип объектов (например, таксофоны, АЗС и т.п.) и одно отношение (например, «быть на расстоянии S от предыдущего пункта»); 5) имеют высокую вычислительную сложность; 6) просты в представлении.

Универсальность ЗК объясняется тем, что она представляет класс NP полных по вычислительной сложности задач в теории алгоритмов. Класс NP (недетерминированные полиномиальные) включает задачи, которые нельзя решить с помощью детерминированной машины Тьюринга за время, растущее с ростом размерности задачи, как полином некоторой фиксированной, не зависящей от размерности задачи, степени. Исследования решения ЗК развиваются в двух направлениях: 1) построение методов решения, позволяющих ускорить ее решение и, соответственно, решать задачи большей размерности; 2) приближение задачи коммивояжера к реальности, добавление объектов и отношений и разработка методов решения таких задач.

Для ускорения поиска оптимального или квазиоптимального пути полезным является использование концепции вычислительных акторов [1], позволяющей эффективно и достаточно просто распараллеливать вычислительные процессы. Вычислительный актер – это модель вычислений, в которой акторы (или процессы) взаимодействуют друг с другом, отправляя и получая сообщения, и могут создавать новых акторов.

Каждый актер может выполнять вычисления, изменять своё состояние и отправлять сообщения другим актерам. Акторы являются независимыми и работают параллельно.

В [14] описано, как использовать модель акторов для построения распределенных систем, в которых акторы могут работать параллельно и взаимодействовать друг с другом через передачу сообщений.

1. Постановка задачи. На плоскости расположены N городов. Требуется найти либо гамильтонов цикл минимальной длины, либо, при невозможности нахождения такого решения за приемлемое время, найти квазиоптимальное решение, отличающееся от оптимального на заранее заданную величину ε . В данном исследовании $\varepsilon=0.01$. Приведем формальную постановку ЗК:

Пусть V – множество вершин графа из N городов, $d_{ij} = d(v_i, v_j)$, $i, j = 1, \dots, N$ – расстояние между парой вершин. Требуется найти минимальное значение $\sum_{i=1}^N d_{\pi(i), \pi(i+1)} + d_{\pi(N), \pi(1)}$ среди всех $(N-1)!/2$ возможных вариантов π , где $\pi(i)$ – i -й город в выбранной перестановке. При этом накладываются условия неотрицательности, симметричности, запрета на петли и удовлетворения неравенству треугольника:

$$d_{ij} \geq 0, d_{ij} = d_{ji}, d_{ii} = \infty, i, j = \overline{1, N},$$

$$d_{ij} + d_{jk} \geq d_{ik}, i, j, k = \overline{1, N} | i \neq j, i \neq k, j \neq k$$

В качестве рабочей гипотезы, исходя из обзора литературы и имеющегося у нас опыта решения подобных задач, мы выбрали следующую совокупность методов, взятых из разных областей знаний, которая в их комплексном применении может дать синергетический эффект для нахождения квазиоптимального решения ЗК.

Поскольку ЗК достаточно просто распараллеливается, в качестве вычислительной архитектуры мы используем концепцию вычислительных акторов [14], параллельно решающих ЗК, обменивающихся частичными результатами, оценивающих, скрещивающих варианты и подвергающих их периодически мутациям в соответствии с МЭС [2]. Хромосомы особей состоят из треугольников триангуляции Делоне, полученной из диаграммы Вороного [4, 15].

2. Диаграмма Вороного и триангуляция Делоне. Следуя [4], введем необходимые определения.

Пусть $P = \{p_1, p_2, \dots, p_n\}$ – множество точек на плоскости, $p_i \in P$ называется сайтом.

Ячейка Вороного ($V(p_i)$) – множество точек плоскости q таких, что для фиксированного сайта p_i и любых других сайтов $p_j \in P, j \neq i$ верно неравенство $\rho(q, p_i) < \rho(q, p_j)$.

Диаграмма Вороного ($Vor(P)$) для сайтов $P = \{p_1, p_2, \dots, p_n\}$ на плоскости – это разбиение плоскости на ячейки Вороного для каждого сайта из P .

Диаграмма Вороного конечного множества точек W на плоскости представляет такое разбиение плоскости, при котором каждая область этого разбиения образует множество точек, более близких к одному из элементов множества W , чем к любому другому его элементу.

Система, обратная диаграмме Вороного, называется триангуляцией Делоне. Эта диаграмма состоит из линий от каждой точки до её ближайших соседей, и каждая линия перпендикулярна пересекаемому ею ребру Вороного.

В [13] приведены алгоритмы триангуляции Делоне для решения ЗК.

Важно отметить, что триангуляция Делоне и диаграммы Вороного являются тесно связанными понятиями в вычислительной геометрии, и их использование при решении ЗК часто включает схожие методы и подходы [16].

Взаимное положение диаграммы Вороного и триангуляции Делоне представлено на рисунке 1. Точками обозначены города – вершины планарного графа, черные линии – границы областей ближайших точек к данному городу, белые линии – ребра треугольников триангуляции Делоне. Белые линии, опирающиеся на стороны квадрата, исключаются из триангуляции Делоне, красные линии образуют границу области триангуляции.

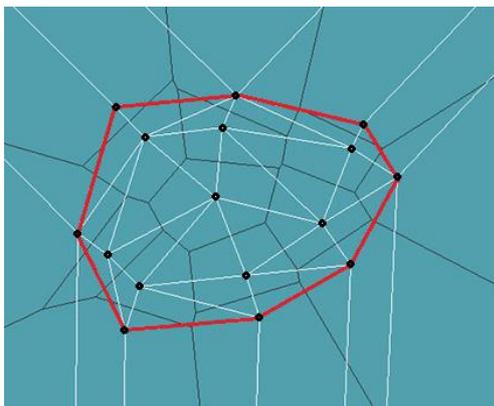


Рис. 1. Диаграмма Вороного и триангуляция Делоне

3. Алгоритм нахождения квазиоптимального пути. Треугольник в триангуляции Делоне представляет собой фигуру, чьи вершины являются городами, а стороны – пути между ними. Существует внешний контур – граница триангуляции.

Примыкающий к границе треугольник – треугольник, у которого одна или две из сторон образуют внешний контур.

Для нахождения гамильтонова цикла будем последовательно ставить метки и удалять из внешнего контура графа треугольники по определенным правилам. После удаления треугольника внешний контур уменьшается.

Правила получения одного из вариантов гамильтонова цикла:

- 1) можно удалить любой треугольник, примыкающий к границе, за исключением тех, с удалением которых удаляется и город;
- 2) можно удалить любой треугольник, примыкающий к удаляемому, за исключением тех, которые могут примкнуть сторонами или углами к другому удаляемому треугольнику или границе;
- 3) процесс удаления треугольников совершается до тех пор, пока на карте не останется городов, не граничащих с удаляемыми треугольниками или внешними границами;
- 4) оставшийся после этих процедур внешний контур и будет гамильтоновым циклом, длина пути у которого, как правило, больше, чем у оптимального решения.

Видно, что для заданного планарного графа число вариантов найденных неоптимальных решений будет существенно меньше, чем у известных к настоящему времени методов поиска квазиоптимальных решений ЗК. Нашей целью является показать, что данный новый метод, обладающий существенно меньшим числом вычислений, может приводить к получению квазиоптимального решения, не превышающего 1.01 длины пути оптимального. Поэтому все наши исследования производились на планарных графах с известным оптимальным путем, и полученные результаты сравнивались. Для этого мы использовали базу данных TSPLIB [17], которая содержит контрольные примеры решения ЗК с проверенными на оптимальность гамильтоновыми циклами. TSPLIB содержит наборы данных ЗК для разных чисел городов N .

Поиск квазиоптимального пути осуществляется с помощью МЭС коллективом вычислительных акторов.

Приведем необходимые определения.

Особь – упорядоченный по неубыванию массив, состоящий из номеров удаленных треугольников.

Популяция – множество особей, число особей равно числу вычислительных акторов K .

Целевая функция – длина пути обхода городов $S = \sum_{j=1}^{M_i} x_i^j$, где M_i – номер треугольника, $i = \overline{1, N}$, x_i^j – длина пути у j -й особи.

Алгоритм поиска квазиоптимального решения ЗК выглядит следующим образом:

На внешнем для вычислительных акторов ресурсе с использованием базы данных TSPLIB формируется гамильтонов цикл минимальной длины для заданного числа городов N и создается таблица расположения в плоскости XOY городов, номера которых взяты из первоначального распределения случайным образом. Длина оптимального пути S запоминается и в дальнейшем используется для оценки качества решения. По методу, описанному в [18], строится диаграмма Вороного и пары городов, лежащих на перпендикуляре к соответствующему ребру диаграммы, образуют ребро треугольника триангуляции Делоне, как это было описано выше. После завершения этой процедуры создается таблица с параметрами треугольников – вершин, ребер и номеров соседних треугольников. Треугольники и их ребра, имеющие общую границу с внешним контуром триангуляции, помечаются.

Вычислительные акторы получают эту таблицу и, выбирая случайным образом начальный треугольник, удаляют соседние треугольники, выбранные также случайным образом в соответствии с правилами, приведенными выше. Процесс продолжается до тех пор, пока не будет получен гамильтонов цикл. Вычисляется длина пути этого цикла. Формируется особь – перечень номеров удаленных треугольников. На этом стадия генерации решений МЭС заканчивается.

На стадии эволюционного согласования решений каждый актор получает особи – варианты решений от двух акторов, номера которых выбраны случайным образом и производит операцию скрещивания. Далее он анализирует полученного потомка на соответствие правилам построения особи и, если находит нарушение какого-либо из правил, то проводит операцию корректировки особи. После определения целевой функции вновь полученной особи – длины пути гамильтонова цикла, актор с заранее заданной вероятностью производит мутацию в одной из случайно выбранных имеющихся у него особей. Для этого он заменяет случайно выбранный номер одного из удаляемых треугольников на другой номер, также полученный случайно. При нарушении правил производится корректировка вновь полученной особи. У подвергнутой мутации особи актор вычисляет целевую функцию. Далее актор оставляет у себя лучшую особь.

После обновления популяции, если получены одинаковые решения – популяция близнецов, или достигнуто заранее заданное число итераций, то квазиоптимальное решение считается найденным и происходит передача лучшего решения внешнему для акторов ресурсу. Если критерии выхода из цикла не достигнуты, то продолжается выполнение стадии эволюционного согласования.

Внешний ресурс из всех присланных вариантов выбирает вариант с лучшей целевой функцией и, сравнивая ее с известной ему оптимальной величиной, определяет, на какую величину получено превышение длины пути квазиоптимального варианта по сравнению с оптимальным.

4. Описание эксперимента. По приведенным выше алгоритмам была написана и отлажена вычислительная программа DELAUNAY. Вычислительные акторы и внешний ресурс смоделированы в ней как отдельные модули, работающие параллельно.

После выборки из базы данных TSPLIB координат городов и запоминания оптимального пути S , происходит передача массива координат вершин следующей части программы – построению диаграммы Вороного и триангуляции Делоне, представленных на рисунке 2.

Далее программа работает в соответствии с алгоритмом нахождения квазиоптимального пути, описанного выше. На выходе программы осуществляется расчет величины превышения квазиоптимального пути над оптимальным. Теоретическая оценка этой величины для МЭС является достаточно трудной задачей и поэтому была проведена статистическая оценка этой величины для $N=100$. При общем количестве испытаний, равном десяти тысячам, это превышение не было больше 0.01, составив среднюю величину в 0.00879. В таблице 1 приведены сравнительные оценки общего количества оцененных вариантов решения ЗК для разных методов – МЭС, полного перебора и метода динамического программирования (МДП).

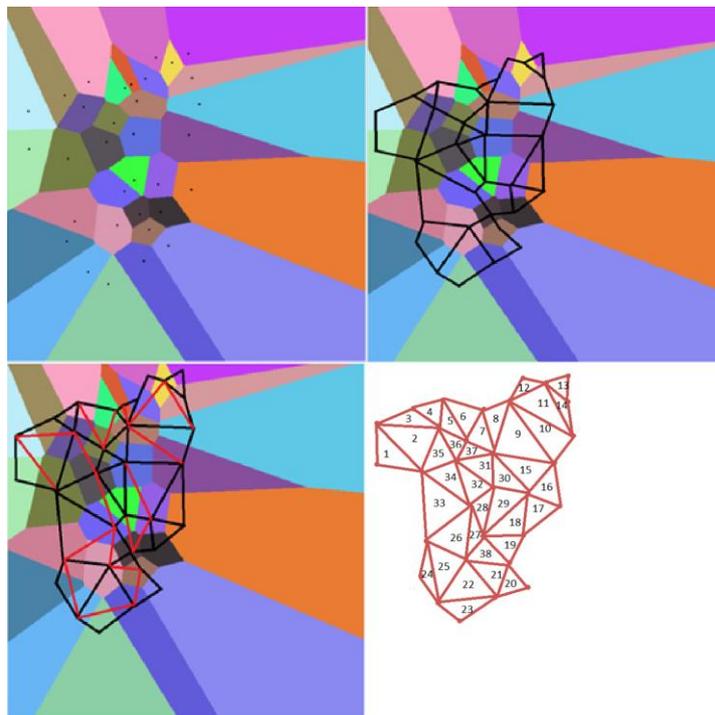


Рис.2. Построение диаграммы Вороного и триангуляции Делоне

Таблица 1. Количество вариантов расчетов для $\varepsilon < 0.01$

Количество городов	МЭС	Полный перебор	МДП
10	155	181440	5120
50	4000	$(50 - 1)!/2$	$50 \cdot 2^{50-1}$
100	15050	$(100 - 1)!/2$	$100 \cdot 2^{100-1}$
200	60100	$(200 - 1)!/2$	$200 \cdot 2^{200-1}$

Статистические оценки работы программы DELAUNAY показали сложность алгоритма $O(n) \sim O(n^2)$.

Заключение. Решение задачи коммивояжера в различных областях применения дает большой экономический эффект, позволяя сокращать время выполнения миссий, расстояния и ресурсы. Применительно к решению этой задачи рассмотрена и исследована совокупность методов, как биоинспирированных, таких, как эволюционное согласование решений и генетические алгоритмы, так и методов вычислительной геометрии в использовании диаграмм Вороного и триангуляций Делоне, а также в целях распараллеливания и ускорения расчетов, использования концепции вычислительных акторов. В сравнении с существующими методами показано преимущество предлагаемого метода, в особенности для среднего и большого

количества вершин планарного графа. Показано, что приведенная совокупность методов, обладая синергетическим эффектом, может быть использована для решения подобных задач. Проведенные вычислительные эксперименты с помощью разработанной программы DELAUNAY показали эффективность такого подхода по сравнению с классическими алгоритмами для задачи коммивояжера. Программа была успешно протестирована на больших входных данных. Была показана приемлемая скорость работы программы.

Предложенная совокупность методов может быть использована и для других задач.

Благодарности. Работа выполнена в рамках государственного задания Минобрнауки России, номер темы FSFF-2023-0005.

Список источников

1. Atkinson R., Hewitt C. Synchronization in actor systems. Proceedings of the 4th ACM SIGACT-SIGPLAN symposium on principles of programming languages, 1977, pp.267-280.
2. Протасов В.И. Конструирование метасистемных переходов / В.И. Протасов. – Протвино: Изд. Института физико-технической информатики, 2009. – 186 с.
3. Lau Sim Rim, Shue Li-Yen Solving travelling salesman problems with an intelligent search approach. Asia-Pacific journal of operational research, 2001, vol.18, pp.77-87.
4. Карабцев С.Н. Построение диаграммы Вороного и определение границ области в методе естественных соседей / С.Н. Карабцев, С.В. Стуколов // Вычислительные технологии, 2008. – Том 13. – №3. – С. 65-80.
5. Володина Е.В. Практическое применение алгоритма решения задачи коммивояжера / Е.В. Володина, Е.А. Студентова // Инженерный вестник Дона, 2015. – №2. – Ч.2. – С. 1-13.
6. Davendra D. Traveling salesman problem, theory and applications. INTECH, 2010, 299 p.
7. Fild M.M. The traveling-salesman problem. Journal of the operations research society of America, 1956, vol.4(1), pp. 61-75.
8. Miller C.E., Tucker A.W., Zemlin R.A. Integer programming formulation of traveling salesman problems. Journal of the ACM, 1960, vol.7(4), pp.326-329.
9. Holland J.H. Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor, 1975, 228 p.
10. Батищев Д.И. Применение генетических алгоритмов к решению задач дискретной оптимизации. Учебное пособие // Д.И. Батищев, Е.А. Неймарк, Н.В. Старостин. – Нижний Новгород: Нижегородский гос. университет им. Н. И. Лобачевского, 2007. – 85 с.
11. Гладков Л.А. Генетические алгоритмы // Л.А. Гладков, В.В. Курейчик, В.М. Курейчик. – М.: Физматлит, 2006. – 367 с.
12. Joseph O'Rourke Computational geometry in C. Cambridge University Press, 1994, 390 p.
13. Liu A. The traveling salesman and delauney triangulation. 2015, available at: <https://web.colby.edu/thegeometricviewpoint/2015/03/09/delauney-triangulations-and-the-traveling-salesman/> (accessed: 10/30/2023).
14. Agha G. Actors – a model of concurrent computation in distributed systems. Computer Science, MIT Press series, 1985, 190 p.
15. Owens B. How to use Voronoi diagrams to control AI, 2013, available at: <https://gamedevelopment.tutsplus.com/how-to-use-voronoi-diagrams-to-control-ai--gamedev-11778t> (accessed: 30.10.2023)
16. Letchford Adam N., Pearson Nicholas A. G4d triangulations yield g4d tours. Computers & operations research 35, 2008, pp. 638-647.
17. База данных TSPLIB. – URL: <http://comopt.ifl.uni-heidelberg.de/software/TSPLIB95/> (дата обращения: 30.10.2023).
18. Препарата Ф. Вычислительная геометрия: Введение / Ф. Препарата, М. Шеймос. — М.: Мир, 1989. – 295 с.

Протасов Владислав Иванович. Доцент, доктор технических наук, Московский авиационный институт (национальный исследовательский университет), профессор кафедры цифровых технологий и информационных систем, AuthorID: 520803, SPIN: 9455-1291, ORCID: 0000-0002-483-7209, protonus@yandex.ru, Россия, Москва, МАИ, Волоколамское шоссе, д. 4.

Потапова Зинаида Евгеньевна. Доцент, кандидат физико-математических наук, Московский авиационный институт (национальный исследовательский университет), доцент кафедры математической кибернетики, AuthorID: 180610, SPIN: 1476-8134, ORCID: 0000-0002-2718-1556, potapovaz@yandex.ru, Россия, Москва, МАИ, Волоколамское шоссе, д. 4.

Мирахмедов Роман Октамович. Аспирант кафедры прикладных программных средств и математических методов Московского авиационного института (национальный исследовательский университет), AuthorID: 1192305, SPIN: 6344-9920, ORCID: 0000-0001-8930, mirakhmedov@gmail.com, Россия, МАИ, Москва, Волоколамское шоссе, д. 4.

Чернова Мария Владиславовна. Студент, Московский авиационный институт (национальный исследовательский университет), стажер кафедры математической кибернетики, lady.keave@gmail.com, Россия, Москва, МАИ, Волоколамское шоссе, д. 4.

UDC 519.854.2

DOI:10.25729/ESI.2023.32.4.004

Quasi-optimal solution of the traveling salesman problem using the evolutionary matching method

Vladislav I. Protasov, Roman O. Mirakhmedov, Zinaida E. Potapova, Maria V. Chernova

Moscow Aviation Institute, Russia, Moscow, *protonus@yandex.ru*

Abstract. An efficient algorithm for the quasi-optimal solution of the traveling salesman problem by a team of actors using the evolutionary matching method is proposed and investigated. The method is based on the use of genetic algorithms. The chromosomes of individuals consist of triangles of the Delaunay triangulation obtained from the Voronoi diagram. The results of the program developed on the basis of the proposed algorithm are compared with existing known methods.

Keywords: Traveling salesman problem, Voronoi diagram, genetic algorithms, Delaunay triangulation, planar graph, evolutionary matching method, computational actors

Acknowledgements: The work was carried out within the framework of the state assignment of the Ministry of Education and Science of Russia, topic number FSFF-2023-0005.

References

1. Atkinson R., Hewitt C. Synchronization in actor systems. Proceedings of the 4th ACM SIGACT-SIGPLAN symposium on Principles of programming languages, 1977, pp.267-280.
2. Protasov V.I. Konstruirovaniye metasisistemnykh perekhodov. The construction of metasytematic transitions. Protvino, Moscow institute of Physics and Technology, 2009, 186 p.
3. Lau Sim Rim, Shue Li-Yen Solving Travelling Salesman Problems with an Intelligent Search Approach, Asia-Pacific Journal of Opereational research, 2001, v.18, pp.77-87.
4. Karabcev S.N., Stukolov S.V. Postroenie diagrammy Voronogo i opredelenie granic oblasti v metode estestvennykh sosedej [Constructing a Voronoi diagram and determining the boundaries of the region in the natural neighbors method]. Vychislitel'nyye tekhnologii [Computing technologies], 2008, no. 13(3), pp. 65-80.
5. Volodina E.V., Studentova E.A. Prakticheskoe primeneniye algoritma resheniya zadachi kommivoyazhera [Practical application of the algorithm for solving the traveling salesman problem]. Inzhenernyy Vestnik Dona [Engineering bulletin of the Don], 2015, no. 2(2), pp. 1-13.
6. Davendra D. Traveling salesman problem, theory and applications. INTECH, 2010, 299 p.
7. Fild M.M. The traveling-salesman problem. Journal of the operations research society of America, 1956, vol.4(1), pp. 61-75.
8. Miller C.E., Tucker A.W., Zemlin R.A. Integer programming formulation of traveling salesman problems. Journal of the ACM, 1960, vol.7(4), pp.326-329.
9. Holland J.H. Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor, 1975, 228 p.
10. Batishchev D.I., Nejmark E.A., Starostin N.V. Primeneniye geneticheskikh algoritmov k resheniyu zadach diskretnoy optimizatsii. Uchebnoe posobie [Application of genetic algorithms to solving discrete optimization problems. Tutorial]. Nizhny Novgorod, Nizhegorodskiy gosudarstvennyy universitet im. N. I. Lobachevskogo [Nizhny Novgorod State University named after N. I. Lobachevsky], 2007, 85 p.
11. Gladkov L.A., Kurejchik V.V., Kurejchik. V.M. Geneticheskie algoritmy [Genetic algorithms]. Moscow, Fizmatlit, 2006, 367 p.
12. Joseph O'Rourke Computational geometry in C. Cambridge University Press, 1994, 390 p.
13. Liu A. The traveling salesman and delauney triangulation. 2015, available at: <https://web.colby.edu/thegeometricviewpoint/2015/03/09/delauney-triangulations-and-the-traveling-salesman/> (accessed: 10/30/2023).

14. Agha G. Actors – a model of concurrent computation in distributed systems. Computer Science, MIT Press series, 1985, 190 p.
15. Owens B. How to use Voronoi diagrams to control AI, 2013, available at: <https://gamedevelopment.tutsplus.com/how-to-use-voronoi-diagrams-to-control-ai--gamedev-11778t> (accessed: 10/30/2023)
16. Letchford Adam N., Pearson Nicholas A. G4d triangulations yield g4d tours. Computers & operations research 35, 2008, pp. 638-647.
17. Baza dannyykh [Database] TSPLIB, available at: <http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/> (accessed: 10/30/2023)
18. Preparata F., Shejmos M. Vychislitel'naya geometriya: Vvedenie [Computational Geometry: Introduction], Moscow, Mir, 1989, 295 p.

Protasov Vladislav Ivanovich. Associate Professor, doctor of technical sciences, Moscow aviation institute, professor of the department of digital technologies and information systems, AuthorID: 520803, SPIN: 9455-1291, ORCID: 0000-0002-483-7209, protonus@yandex.ru, MAI, Russia, Moscow, Volokolamskoye Shosse, 4.

Potapova Zinaida Evgenievna. Associate professor, Candidate of physical and mathematical sciences, Moscow aviation institute, associate professor of the department of mathematical cybernetics, AuthorID: 180610, SPIN: 1476-8134, ORCID 0000-0002-2718-1556, potapovaz@yandex.ru, 125993, Russia, Moscow, MAI, Volokolamskoye Shosse, 4.

Mirakhmedov Roman Oktamovich. Postgraduate student, Moscow aviation institute, postgraduate student of the department of applied software and mathematical methods, AuthorID: 1192305, SPIN: 6344-9920, ORCID 0000-0001-8930, mirakhmedov@gmail.com, 125993, Russia, Moscow, MAI, Volokolamskoye Shosse, 4.

Chernova Maria Vladislavovna. Student, Moscow aviation institute, intern at the department of mathematical cybernetics, lady.keave@gmail.com, 125993, Russia, Moscow, MAI Volokolamskoye Shosse, 4.

Статья поступила в редакцию 17.10.2023; одобрена после рецензирования 30.10.2023; принята к публикации 13.12.2023.

The article was submitted 10/17/2023; approved after reviewing 10/30/2023; accepted for publication 12/13/2023.

УДК 303.094

DOI:10.25729/ESI.2023.32.4.005

Логика множеств и логика волн в когнитивно-поведенческом моделировании

Суров Илья Алексеевич

Университет ИТМО, Россия, Санкт-Петербург, ilya.a.surov@itmo.ru

Аннотация. Когнитивно-поведенческие процессы часто выходят за рамки классических моделей, предполагающих рациональную алгоритмичность человека. В статье предпринят пересмотр логических установок, обуславливающих эти ограничения. Показано, что одной из таких установок является использование классической теории вероятностей и логики множеств (в том числе нечётких), иногда ошибочно рассматриваемых как безальтернативные. Соответственно общеприродному дуализму дискретность – непрерывность и частица – волна, такой альтернативой является волновая логика и соответствующее ей вероятностное исчисление. Эффективность этой логики в поведенческом моделировании показана на примере экспериментов «дилемма заключённого» и «двухэтапная игра». Волновая вероятностная модель добавляет к логике множеств дополнительный интерференционный фактор, предсказывающий наблюдаемые в этих экспериментах отклонения от классической рациональности. Определяющие этот фактор фазы когнитивных волн кодируют субъективно-смысловые закономерности естественного мышления, не учитываемые в классических моделях. Эти новые закономерности позволяют использовать волновую модель для прогнозирования вероятностей принятия «иррациональных» решений, в том числе, в новых экспериментальных условиях. Рассмотренные свойства логики волн, общие для оптических, голографических и квантовых подходов к обработке информации, открывают новые возможности в задачах когнитивно-поведенческого моделирования и анализа данных.

Ключевые слова: волновая логика, когнитивная интерференция, поведенческое моделирование, вероятностное прогнозирование, теория множеств, рациональность

Цитирование: Суров И.А. Логика множеств и логика волн в когнитивно-поведенческом моделировании / И.А. Суров // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 51-66. – DOI:10.25729/ESI.2023.32.4.005.

Введение. Управление социально-экономическими системами опирается на модели, прогнозирующие когнитивно-поведенческую динамику действующих лиц. Трудность такого моделирования состоит в том, что, в отличие от механики инертных и детерминированных систем (камень, робот, электростанция без операторов), течение этих процессов предопределено законами физики не полностью. Их прогнозирование требует привлечения субъективно-смысловых категорий психологии, социологии и других гуманитарных наук. Такие описания, однако, обычно не имеют математического выражения, что затрудняет их использование на практике. В этой связи встаёт вопрос о пересмотре фундаментальных предпосылок используемой при этом математики, обуславливающих ограниченность имеющихся количественных моделей.

Одной из таких предпосылок является выбор используемой логики. По умолчанию обычно применяется так называемая булева логика, согласно которой любое событие обязано быть либо истинным, либо ложным [1]. Эта логика и выражающая её алгебра множеств лежат в основе классической теории вероятности и рационального экономического поведения [2, 3]. Социально-экономические кризисы последних десятилетий указывают на усугубляющуюся неадекватность этих теорий, не позволяющих предвидеть и управлять течением общественно-важных процессов. Выясняется, что красивое здание социально-экономических наук во многом основано на мифах, имеющих с реальностью мало общего [4–7].

Для решения этих проблем разработан ряд теоретических подходов на волновых, голографических и квантовых принципах. Удалось построить, в частности, количественные модели принятия иррациональных решений [8, 9], неклассической динамики рыночных цен [10, 11], лавинных информационно-поведенческих процессов [12, 13] и феноменальной устойчивости биологических, экологических и социальных систем [14, 15]. Исследуются механизмы

информационно-полевого взаимодействия людей и сообществ [16–20], разрабатываются новые парадигмы субъективно-денежной ценности [21, 22] и общественных отношений [23, 24]. Систематическое развитие этих подходов, однако, затруднено отсутствием общего логического основания, которое бы проявило причины получаемых преимуществ и позволило использовать их более эффективно.

В статье представлен вариант такого основания в виде логики волновых процессов или *логики волн*. Принципы этой логики показаны на примере поведенческих экспериментов «двухэтапная игра» и «дилемма заключённого», описанных в разделе 1. Описание этих данных в рамках логики множеств и классической теории вероятности приведено в разделе 2. Вероятностная модель на основе волновой логики и её апробация в описательном и предсказательном режимах представлены в разделах 3 и 4. В разделе 5 обсуждается материальный носитель когнитивных волн, а также рассмотрены информационно-смысловые свойства волновой логики.

1. Экспериментальные данные. В качестве примера далее использованы эксперименты «двухэтапная игра» и «дилемма заключённого» [25], хорошо известные в когнитивных исследованиях. В каждом из них целевое двухвариантное решение типа «делать – не делать» принимается в трёх различных ситуациях – контекстах.

1.1 Двухэтапная игра. В двухэтапной игре испытуемый принимает решение участвовать или не участвовать в следующем коне игры в орлянку, задача которой состоит в угадывании стороны подброшенной монеты. В случае выигрыша испытуемый получает 200 условных единиц, а в случае проигрыша теряет 100 единиц. В рассматриваемом случае «честной» игры вероятности выигрыша и проигрыша одинаковы, так что математическое ожидание выигрыша составляет $0,5 * 200 + 0,5 * (-100) = 50$. Положительность этой величины указывает на то, что рациональным решением является участие в игре независимо от предыстории.

На практике это ожидание не выполняется и решение об участии (A) рассматривается в трёх ситуациях: когда предыдущий кон выигран (B_1), когда предыдущий кон проигран (B_2), и когда исход предыдущего кона неизвестен (B_1 или B_2). В каждом случае решение принимается выборкой испытуемых, на основе чего вычисляются значения вероятностей $p(A|B_1)$, $p(A|B_2)$ и $p(A)$ соответственно. Эти величины, полученные в 12 постановках эксперимента, показаны в левой части таблицы 1. Постановки 1-3: [26], 4-7: [27], 8-10: [28], 11: [29]. 12: среднее по экспериментам в работе [30]. Как отмечено выше, во всех случаях решения испытуемых отклоняются от рационального ожидания $p(A|B_1) = p(A|B_2) = p(A) = 1$.

1.2 Дилемма заключённого. В «дилемме заключённого» испытуемый и его напарник находятся под следствием. Испытуемому предлагается принять решение о предательстве напарника (A) или молчании в трёх ситуациях: когда напарник молчит (B_1); когда напарник его предал (B_2); когда поведение напарника неизвестно (B_1 или B_2). При этом испытуемому сообщается, что если оба подсудимых молчат, то они получают сравнительно небольшое наказание, например, в виде 1 года тюремного заключения; если оба предали, то оба получают по 3 года; если же один предал, а второй промолчал, то первый выходит на свободу, а второй садится на 10 лет.

Согласно этим правилам, оптимальным решением в любом случае является предательство, так что рациональное поведение соответствует вероятностям $p(A|B_1) = p(A|B_2) = p(A) = 1$, как и в двухэтапной игре. Результаты 12 постановок этого эксперимента представлены в правой части таблицы 1. Постановка 1: [25], 2-3 [31], 4: [32], 5: [33], 6-10: [34], 11-12: [35]. Во всех случаях решения отклоняются от рационального ожидания.

Таблица 1. Результаты 12 постановок экспериментов «двухэтапная игра» и «дилемма заключённого»

№	Двухэтапная игра					Дилемма заключённого				
	$p(A B_1)$	$p(A B_2)$	$p(A)$	Δ	$\delta,^\circ$	$p(A B_1)$	$p(A B_2)$	$p(A)$	Δ	$\delta,^\circ$
1	0,69	0,57	0,38	0,25	113,5	0,84	0,97	0,63	0,28	107,7
2	0,75	0,69	0,73	-0,01	89,2	0,17	0,68	0,22	0,2	126,0
3	0,69	0,59	0,35	0,29	117,0	0,47	0,65	0,38	0,18	109,0
4	0,60	0,47	0,47	0,07	97,3	0,84	0,91	0,66	0,21	104,2
5	0,83	0,70	0,62	0,14	101,1	0,93	0,97	0,88	0,07	94,2
6	0,80	0,37	0,43	0,16	106,7	0,67	0,73	0,6	0,1	98,2
7	0,68	0,32	0,38	0,12	105,4	0,77	0,8	0,63	0,16	101,4
8	0,64	0,47	0,38	0,18	108,6	0,87	0,9	0,87	0,02	91,0
9	0,53	0,38	0,24	0,22	118,6	0,8	0,83	0,64	0,18	102,4
10	0,73	0,49	0,60	0,01	91,0	0,73	0,83	0,77	0,01	90,7
11	0,30	0,24	0,17	0,1	111,9	0,87	0,94	0,85	0,05	93,2
12	0,59	0,71	0,70	-0,05	85,7	0,89	0,93	0,86	0,06	93,5
Ср.	0,65	0,50	0,45	0,12	104	0,74	0,85	0,67	0,13	101

2. Классическая вероятность и логика множеств.

2.1 Элементарная модель. Классическая вероятностная модель для эксперимента «двухэтапная игра» (разд. 1.1) строится на основе теоретико-множественного представления событий. Следуя аксиоматике А.Н. Колмогорова [2], отождествим элементарное событие с пространственным положением всех задействованных в эксперименте предметов и тел, включая монету и речевой аппарат игрока. Полное множество Ω таких элементарных событий показано на рис. 1А прямоугольником, площадь которого соответствует вероятностной мере единица. Исходам предыдущего кона «выигрыш» B_1 и «проигрыш» B_2 соответствуют подмножества Ω , охватывающие соответствующие им положения монеты и ответы игрока. Эти исходы являются взаимоисключающими, так что вероятностная мера их одновременного наблюдения, то есть площадь пересечения соответствующих частей прямоугольника, равна нулю: $p(B_1 \cap B_2) = 0$. Полагая другие исходы эксперимента («на ребро» и т.д.) невероятными, эти два сложных события полностью охватывают множество элементарных событий Ω , так что вероятность объединения этих событий, то есть сумма их площадей, равна единице: $p(B_1 \cup B_2) = p(\Omega) = 1$. Соответственно, вероятности условий B_1 и B_2 подчиняются соотношению:

$$p(B_1) = 1 - p(B_2) = x, \quad 0 \leq x \leq 1, \tag{1}$$

где параметр x – априорная вероятность условия B_1 – введён для краткости дальнейших формул.

Аналогично событиям B_1 и B_2 , решение A играть в следующем коне представлено на рис. 1 областью, частично пересекающейся с альтернативными условиями B_1 и B_2 ¹. Соответственно, вероятность этого решения $p(A) = p(A|B_1 \cup B_2)$ должна равняться сумме площадей этих пересечений:

$$p(A) = p(A \cap B_1) + p(A \cap B_2), \tag{2}$$

что соответствует распределительному закону, то есть свойству дистрибутивности булевой логики [36].

Согласно формуле Байеса, слагаемые в этом выражении представляются в виде произведения условных и априорных вероятностей:

¹ При этом предполагается, что с момента завершения предыдущего кона результатом B_1 или B_2 множество элементарных событий Ω и заданная на нём вероятностная мера не изменились.

$$p(A \cap B_k) = p(A|B_k) * p(B_k), \quad k \in \{1,2\} \quad (3)$$

так что формула (2) принимает вид

$$p(A) = \sum_{k=1,2} p(A|B_k)p(B_k) = p(A|B_1) * x + p(A|B_2) * (1 - x), \quad (4)$$

известный как закон полной вероятности: вероятность события A есть сумма его вероятностей при всех возможных условиях, помноженных на априорные вероятности этих условий [36]. В случае «честной» игры $x = 0,5$ формула (4) принимает вид:

$$p(A) = \frac{p(A|B_1) + p(A|B_2)}{2}, \quad (5)$$

согласно которому вероятность $p(A)$ игры при неизвестном исходе есть среднее от вероятностей игры при выигрыше и проигрыше предыдущего кона. В общем случае (4) вероятность $p(A)$ может лежать в любой точке интервала от $p(A|B_1)$ до $p(A|B_2)$ в зависимости от параметра x (1).

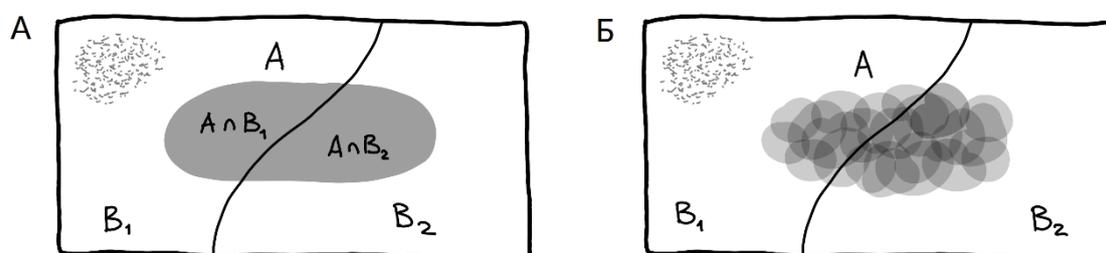


Рис. 1. Классическая вероятностная модель двухэтапной игры и дилеммы заключённого (разд. 1), построенная на основе логики чётких (А) и нечётких (Б) множеств

2.2. Нечёткая логика. Для приближения представленной модели к логике естественного мышления используется аппарат нечётких множеств, согласно которому принадлежность элементарного события x к сложному событию A описывается не дискретными значениями 1 или 0, а непрерывной функцией принадлежности $\mu_A(x)$ [37, 38]. Соответственно, «нечёткое» событие A на множестве Ω изображается не резко очерченной фигурой как показано на рис. 1А, а градиентным облаком как на рис. 1Б. Вероятность $p(A)$ тогда можно отождествить с интегралом функции $\mu_A(x)$ по множеству событий Ω .

Поскольку условия B_1 и B_2 по-прежнему являются взаимоисключающими, граница между соответствующими областями пространства событий на рис. 1Б остаётся такой же чёткой, как и на рис. 1А. В том же предположении статичности пространства событий (сноска 1) совместные вероятности $p(A \cap B_k)$ в правой части (2) тогда равны интегралам функции принадлежности $\mu_A(x)$ по областям B_k пространства Ω . Поскольку области B_1 и B_2 не пересекаются, эти величины по-прежнему связаны с условными вероятностями $p(A|B_k)$ правилом Байеса (3), в силу чего соотношения (4) и (5) сохраняют свою силу.

2.3. Проблема. Предсказания (4) и (5) логики множеств проверяемы на основе экспериментальных данных, представленных в разделе 1. Для первой постановки эксперимента по двухэтапной игре (первая строка таблицы 1) среднее от условных вероятностей $p(A|B_1) = 0,69$ и $p(A|B_2) = 0,57$ есть 0,63, что согласно (5) должно совпадать с вероятностью $p(A)$ игры при неизвестном исходе предыдущего кона. Фактически измеренное значение этой величины, однако, составило 0,38 при отклонении от классического ожидания $\Delta = 0,63 - 0,38 = 0,25$. Как видно из таблицы 1, подобное расхождение имеет место для остальных постановок. Замечательно, что при этом измеренная вероятность $p(A)$ не только отличается от среднего по

условным вероятностям, но меньше наименьшей из них. Этот результат нельзя объяснить даже более общим выражением (4), полагая, например, что субъективно-воспринимаемая вероятность выигрыша x отличается от заданной в условии $x = 0,5$.

Аналогичное нарушение имеет место в эксперименте «дилемма заключённого», постановки которого представлены в правой части таблицы 1. Во всех случаях разность Δ между правой и левой сторонами (5) отлична от нуля. Эти результаты статистически достоверно указывают на непригодность вероятностной модели, построенной в разделе 2.1. Нарушения такого типа имеют место не только в обыденно-бытовых решениях, но и в экспертной аналитике экономических, экологических и политических процессов [25, 39]. При этом нарушается не только дистрибутивность, но и коммутативность, монотонность, закон исключённого третьего и другие свойства логики множеств [9].

Эти результаты, однако, не означают, что объяснение данных рассмотренного типа в рамках классической теории вероятности и байесовского вывода в принципе невозможно. Такое объяснение возможно на основе моделей с большим числом параметров, определяющих логику принятия решений аналогично, например, теории перспектив [40]. Дополнительные степени свободы делают такие модели более гибкими, однако одновременно с этим снижают их интерпретируемость и теоретическую ценность. Модель [41], например, позволяет описать 14 типов «иррационального» поведения с помощью 11 подгоночных параметров, определяемых на основе тренировочного массива данных. В пределе такая логика ведёт к моделям машинного обучения, кодирующим логику принятия решений внутренним состоянием нейронной сети. Интерпретация «чёрных ящиков» такого типа на сегодняшний день является нерешённой задачей, а их теоретическая ценность близка к нулю.

Эти и другие проблемы обусловили необходимость пересмотра логико-множественных оснований современных моделей естественного и искусственного мышления [42, 43]. Результаты такого пересмотра осмыслены в работах [44, 45] где установлено, что критическим недостатком таких моделей является статичность событийного пространства Ω . В представленной модели место этого предположения отмечено сноской 1. Фактически, предположение статичности исключает возможность появления в ходе рассматриваемого процесса новых возможностей, в результате чего его модельное течение сводится к выборке из фиксированного множества Ω различных комбинаций элементарных событий². Нарушения этого предположения и строящихся на его основе вероятностных моделей обусловлены творческим (неалгоритмическим) поведением живых организмов, участвующих в процессе.

3. Волновая логика и вероятностная модель. Естественные основания и практические ограничения логических моделей хорошо изучены в физике, где логика множеств выражается в атомистической парадигме. Элементами физических множеств являются, например, молекулы и атомы в составе твёрдых тел, жидкостей и газов. При этом явления природы объясняются более или менее сложным движением частиц в пространстве согласно законам Ньютона. Такая логика частиц позволила получить математические модели классической механики, термодинамики и физики сплошных сред.

Для ряда явлений оптики, электромагнетизма и квантовой механики, однако, такой подход оказался неприменим, в связи с чем для них разработаны различные разделы физики *волн*. Модели соответствующих явлений строятся на основе понятия о поле – непрерывной сущности, обладающей энергией и несущей информацию³. В отличие от плотного вещества, физические поля (электрические, магнитные, гравитационные и другие) не состоят из частиц, в

² Предположение статичности обусловлено фундаментальной установкой абсолютной предопределённости (детерминизма), характерной для механистического мировоззрения [63].

³ В обыденном языке полю соответствует понятие «дух», обозначающее невидимые силы и свойства, присущие живой природе [64, с. 518].

связи с чем логика множеств для их описания не подходит. События в полевых процессах подчиняются принципиально другой, волновой логике или *логике волн*, также нашедшей математическое выражение. Для целей этой статьи достаточно рассмотреть следующую простейшую модель.

3.1 Интерференция волн. Физическим аналогом рассматриваемых экспериментов является классический эксперимент по изучению волновой природы света, известный как двухщелевая интерференция Т. Юнга. В этом эксперименте (1803 год) источник света освещает экран по другую сторону от препятствия с двумя щелевидными апертурами, как показано на рис. 2. В логике множеств интенсивность света на экране определяется потоком приходящих частиц, который равен сумме потоков, приходящих от каждой щели. На практике, однако, вместо наложения двух засветок наблюдается периодическое распределение интенсивности, схематично показанное в правой части рисунка. При этом на экране появляются области, где интенсивность меньше каждой из складывающихся компонент, а также больше их суммы, что нельзя объяснить в рамках частичной (корпускулярной) теории света. Это явление аналогично выходу поведенческих вероятностей $p(A)$ за рамки интервала между $p(A|B_1)$ и $p(A|B_2)$.

Модель, описывающая поведение света в эксперименте Юнга, строится так. Положим, что источник света порождает расходящуюся *волну* возмущения некоторой среды (воздух, вода, «светоносный эфир», «физический вакуум»), амплитуда которой зависит от времени по гармоническому закону $E \cos(\omega t)$. Согласно принципу Гюйгенса-Френеля [46, гл. 4], поле на экране складывается из частей исходной волны, прошедших через щели в препятствии. В рассматриваемую точку на экране эти вторичные волны исходной частоты ω прибывают с амплитудами:

$$\begin{aligned} A_1 &= \sqrt{2x} * E_1 \cos(\omega t + \phi_1) \\ A_2 &= \sqrt{2(1-x)} * E_2 \cos(\omega t + \phi_2), \end{aligned} \tag{6}$$

зависящими от набегов фаз ϕ_k на пройденных путях и соответствующих ослаблений исходной амплитуды E . В этом ослаблении параметр $0 \leq x \leq 1$ учитывает относительное открытие щелей, осуществляемое, например, вертикальным смещением средней части препятствия, как показано слева на рис. 2.

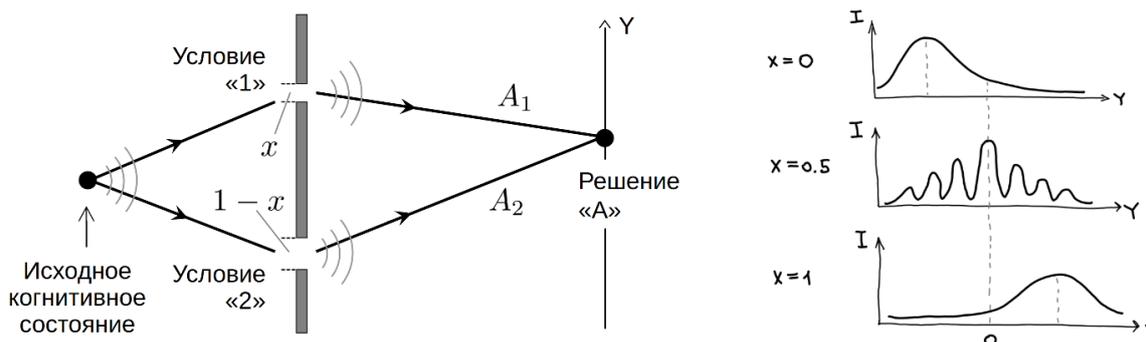


Рис. 2. Двухщелевая интерференция света как физический прототип волновой когнитивно-поведенческой модели. Источнику излучения соответствует исходное когнитивное состояние субъекта. Условиям B_k соответствуют альтернативные пути распространения (когнитивных) волн. Целевое решение «А» соответствует некоторой точке на экране Y . Справа: *интенсивность сигнала на экране при различных положениях препятствия x*

Если открыта только одна из щелей ($x = 1$ или $x = 0$), то интенсивность света в рассматриваемой точке определяется (усреднённым по времени) *квадратом* соответствующей амплитуды (6):

$$I_k = \langle A_k^2 \rangle = 2E_k^2 \left\langle \frac{1 + \cos 2(\omega t + \phi_k)}{2} \right\rangle = E_k^2. \quad (7)$$

Если же открыты обе щели, то интенсивность света в той же точке порождается общей волной суммарной амплитуды, как:

$$I = \langle (A_1 + A_2)^2 \rangle = xE_1^2 + (1-x)E_2^2 + 2\sqrt{x(1-x)}E_1E_2 \cos(\phi_2 - \phi_1). \quad (8)$$

Первые два слагаемых в правой части (8) соответствуют интенсивностям отдельных источников (7), взвешенных относительными ширинами щелей. Третье слагаемое есть так называемый *интерференционный член*, характерный для логики волн.

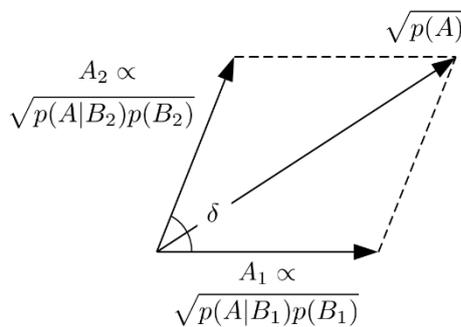
Форма выражения (8) позволяет изобразить его в графическом виде, как показано на рис. 3А, 3Б. Согласно этой схеме амплитуды источников (5) без временной зависимости складываются векторным образом, в результате чего модуль суммарного вектора зависит от угла между слагаемыми. Этот угол $\delta = \phi_2 - \phi_1$, равный разности фаз интерферирующих волн, является отличительным параметром волновой модели.

Штриховками обозначены области конструктивной и деструктивной интерференции с экстремумами в указанных точках. Предсказание классической вероятностной модели (4) показано наклонной прямой

3.2 Когнитивно-поведенческая формулировка. Как отмечено выше, отклонение интенсивности света на экране от предсказания корпускулярной теории аналогично выходу вероятностей $p(A)$ за рамки интервала между $p(A|B_1)$ и $p(A|B_2)$, рассмотренному в разделе 2.3. Это сходство позволяет объяснить когнитивно-поведенческие нарушения логики множеств на основе волновой оптики. Для этого элементы представленной физической модели соотносятся с элементами экспериментов по двухэтапной игре (раздел 1.1) и дилемме заключённого (раздел 1.2) следующим образом.

Источнику света соответствует исходное психическое, или когнитивное состояние человека, представляемое точкой в некотором психическом пространстве. Принятию целевого решения в этом пространстве соответствует переход из исходной точки в точку А на экране, как показано на рис. 2. Аналогично распространению света, этот переход имеет характер волнового процесса, который назовём *когнитивной волной*. Как и в физике, эта волна может достигнуть точки А по двум различным путям, соответствующим различным состояниям промежуточного условия B_k .

А



Б

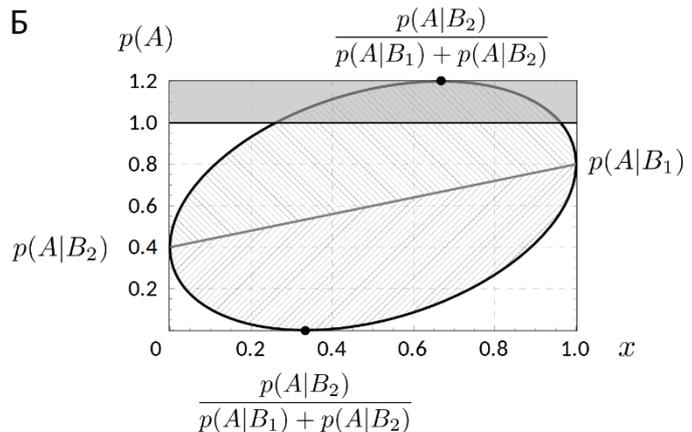


Рис. 3. А: Векторная диаграмма интерференции когнитивных волн (6), (8), порождающей волновую вероятностную модель (10), (11).

Рис. 3. Б: Диапазон значений вероятности $p(A)$ при фиксированных $p(A|B_1) = 0.8$ и $p(A|B_2) = 0.4$, объяснимый с помощью волновой модели (10), в зависимости от вероятности $x = p(B_1)$.

Вероятностям принятия решений во всех случаях соответствует интенсивность сигнала в точке А, также принимающая (в отличие от амплитуд когнитивных волн) неотрицательные значения, а именно, интенсивности света от каждой из щелей по отдельности соответствуют условным вероятностям решения А при условиях B_1 и B_2 , тогда как вероятность решения А при неизвестном условии соответствует интенсивности от обеих щелей одновременно:

$$I_k = E_k^2 = p(A|B_k), \quad I = p(A). \quad (9)$$

Априорные вероятности условий B_k определяются параметром открытия щелей x (рис. 2), определяющим относительный вес альтернативных траекторий в когнитивном пространстве согласно формуле (1). Посредством замен (1) и (9) выражение для интенсивности света (8) переходит в соотношение поведенческих вероятностей:

$$p(A) = \sum_{k=1,2} p(A|B_k)p(B_k) + 2 \sqrt{\prod_{k=1,2} p(A|B_k)p(B_k)} \cos \delta. \quad (10)$$

В простейшем и основном случае $p(B_1) = p(B_2) = 0,5$ это выражение принимает вид:

$$p(A) = \frac{p(A|B_1) + p(A|B_2)}{2} + \sqrt{p(A|B_1) * p(A|B_2)} \cos \delta. \quad (11)$$

Первым слагаемым в правых частях (10) и (11) являются предсказания классической вероятностной модели (4) и (5), которые обычно отождествляются с рациональной логикой принятия решений [47]. Соответственно, вторые слагаемые в правых частях (10) и (11) позволяют описать *иррациональную* компоненту естественного мышления. Эта компонента, иногда называемая фактором привлекательности [48], и отличает логику волн от логики множеств.

Ключевым параметром волновой модели является разность фаз δ , определяющая соотношение вероятностей целевого решения $p(A)$, $p(A|B_1)$ и $p(A|B_2)$ в различных условиях. В симметричном случае (11) он вычисляется, как:

$$\cos \delta = \frac{2p(A) - p(A|B_1) - p(A|B_2)}{2\sqrt{p(A|B_1) * p(A|B_2)}}, \quad (12)$$

что завершает построение волновой модели в описательном (подгоночном) режиме.

Величина отклонения поведенческих вероятностей от классических прогнозов (4) и (5), которую можно описать волновой моделью, ограничивается максимальной величиной второго слагаемого в (10) и (11). В формуле (12) это соответствует равенству $\cos \delta = \pm 1$. В общем случае (10) допустимый диапазон вероятности $p(A)$ при фиксированных $p(A|B_1) = 0,8$ и $p(A|B_2) = 0,4$ показан на рис. 3Б в зависимости от параметра x (1). За границами этого диапазона (белая область за границами штриховки) решения уравнений (10), (12) в действительных числах не существует, что ограничивает степень иррациональности, описываемую в рамках представленной модели⁴.

4. Результаты

4.1 Описательное моделирование. Представленная в разд. 3 волновая модель (10), (11) построена для каждой постановки двухэтапной игры и дилеммы заключённого. Согласно условиям экспериментов (разд. 1), в обоих случаях использовалось значение $x = 0,5$, соответствующее равновероятному осуществлению условий B_1 и B_2 . При этом уравнение (12) решается в действительных числах для всех постановок. Полученные значения фазового параметра δ приведены в таблице 1.

⁴ Для снятия этого ограничения рассматривается возможность так называемой гиперболической интерференции [65], физические предпосылки для которой, однако, не ясны.

В большинстве постановок параметр δ превышает 90° , что соответствует отрицательному значению интерференционного члена в (11). Это означает, что фактическое значение вероятности $p(A)$ меньше рационального ожидания (4), соответствующего на рис. 3Б середине серой прямой. Для произвольных значений x этому режиму «деструктивной» интерференции соответствует нижняя половина эллипса на рис. 3Б, в которой амплитуды когнитивных волн (рис. 2А) складываются в противофазе и частично подавляют друг друга. На векторной диаграмме (рис. 2Б) это соответствует тупому углу между амплитудами A_1 и A_2 . Для каждой постановки обоих экспериментов такие диаграммы показаны на рис. 4.

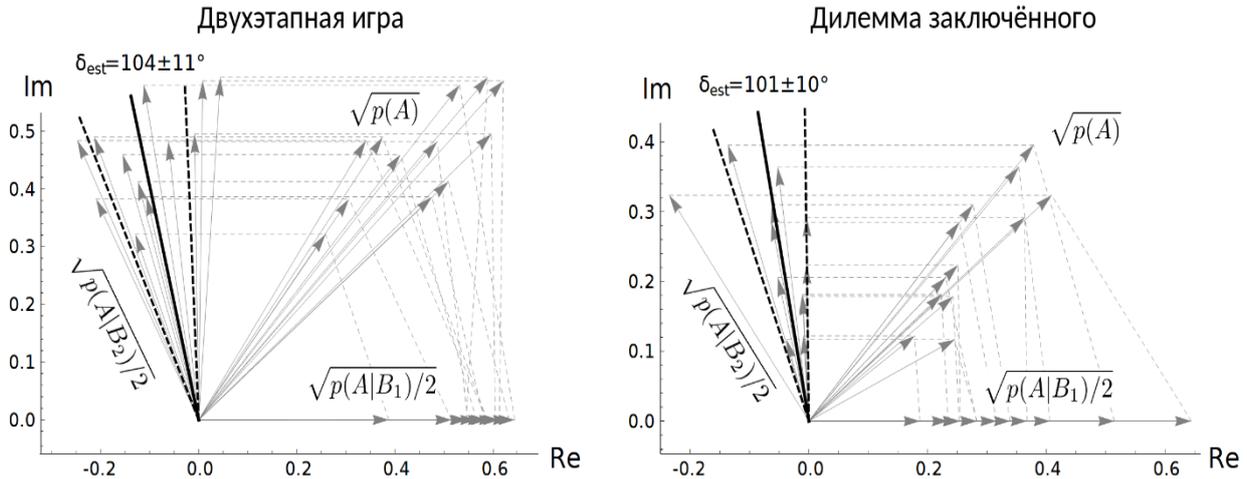


Рис. 4. Векторные диаграммы (рис. 3А) волновой модели (11) для каждой постановки экспериментов по двухэтапной игре и дилемме заключённого согласно таблице 1

4.2 Фазовая стабильность и прогнозное моделирование. Векторные диаграммы на рис. 4 характерны угловой группировкой амплитуд $\sqrt{p(A|B_2)}/2$. Для двухэтапной игры и дилеммы заключённого соответствующие разбросы составляют $104 \pm 11^\circ$ и $101 \pm 10^\circ$ соответственно. Эта устойчивость фазового параметра δ может быть экстраполирована на новые эксперименты такого же типа. Полагая величину δ лежащей в указанном диапазоне, формула (11) тогда позволяет рассчитать вероятность $p(A)$ на основе значений $p(A|B_1)$ и $p(A|B_2)$ для новых постановок двухэтапной игры. Возможность такого прогноза проверена в эксперименте [29], для которого условные вероятности составили $p(A|B_1) = 0,30$ и $p(A|B_2) = 0,24$. Спрогнозированный интервал $p(A)_{\text{прог}} = 0,19 \pm 0,03$ накрыл фактически измеренное значение $p(A)_{\text{факт}} = 0,17$. В работе [49] сходная точность получена для других экспериментов подобного типа.

Данные, представленные в таблице 1, позволяют проверить эту методику прогнозного моделирования более полно. Для обоих экспериментов использовался следующий алгоритм:

1. Выбрать одну из 12 постановок эксперимента;
2. Взять среднее значение фазового параметра δ для остальных 11 постановок;
3. Использовать его для предсказания $p(A)$ в выбранной постановке, как показано выше.

Таким образом, исход каждой постановки эксперимента «прогнозируется» на основе остальных 11 постановок того же самого эксперимента. На рис. 5 полученные прогнозные значения показаны в сравнении с фактическими величинами в таблице 1.

Для двухэтапной игры среднеквадратичное отклонение прогноза от факта $p(A)$ составило 0,12, то есть 26% от средней величины $p(A)$. Для дилеммы заключённого эти величины составили 0,11 и 33% соответственно. Полученная точность близка к точности линейной ре-

грессии с тремя свободными параметрами, по результатам которой среднеквадратическое отклонение прогноза от факта составило 0,12 для обоих экспериментов. Волновая модель, однако, существенно превосходит линейную регрессию по доле объяснённой дисперсии (коэффициенту детерминации R^2). Для двухэтапной игры и дилеммы заключённого R^2 возрастает с 0,50 до 0,55 и с 0,52 до 0,79 соответственно.

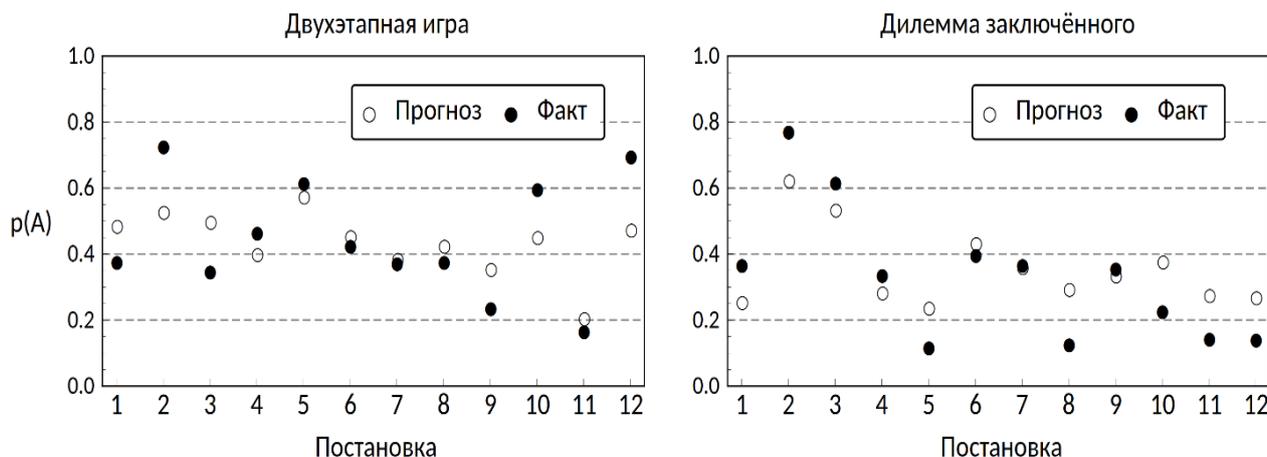


Рис. 5. Прогнозирование иррациональных решений на основе устойчивых величин интерференционной фазы δ волновой вероятностной модели. Прогнозные и фактические значения вероятности целевого решения «А» для постановок двухэтапной игры (слева) и дилеммы заключённого (справа) показаны белыми и чёрными точками

5 Носитель и семантика логики волн

5.1 Носитель когнитивных волн. Как показано в разделе 4, вероятностные соотношения (10) – (12) могут использоваться для поведенческого моделирования формально, не уточняя материальный носитель когнитивных волн и механизм их интерференции. Однозначных ответов на эти вопросы в настоящее время нет. В качестве гипотезы рассматриваются электромагнитные колебания в нервной системе организмов, включая волновую динамику нейронных ансамблей и потенциалов действия [49–51]. Эта гипотеза сопрягает волновую логику с многочисленными моделями нейрофизиологии, среди которых особо интересны фазово-чувствительные [52–55].

Знание биологических носителей, тем не менее, желательно для дальнейшего развития представленной модели. От этих носителей зависит, например, время усреднения в (7) и (8), в этих выражениях устремлённое к бесконечности. Кроме того, реальные волны ограниченного пространственного размера всегда имеют ненулевую спектральную ширину и конечную длину когерентности. Учёт этих эффектов ведёт к частичному подавлению интерференционных членов в выражениях (8), (10) и (11), приближая их к классическим аналогам (4) и (5); то же самое имеет место при расхождении частот интерферирующих волн. В результате появляется возможность моделирования логик с переменным соотношением волнового и частичного аспектов, тогда как логика множеств и логика волн становятся предельными случаями этой обобщённой модели аналогично соотношению волновой оптики и её геометрического предела, а также квантовой и классической механики [56]. Эти и другие биофизические параллели подсказывают перспективные направления развития волновой логики на принципах естественного мышления. Сходным образом возможно использование моделей электромагнетизма, акустики, волновой и квантовой оптики, лазерной физики и взаимодействия излучения с веществом [12, 57, 58].

5.2 Семантика логики волн. Волна есть одна из наиболее простых и распространённых форм движения множества частиц. Без частицы нет волны, а у любого состояния частиц есть

форма. Таким образом дуализм логики множеств и логики волн, по существу, аналогичен дуализму материи и информации в естественном мышлении и природе в целом [59, 60]. То же самое верно для более общего формализма, в котором место волн занимают квантовые состояния [29]. И когнитивные волны, и квантово-когнитивные состояния есть единицы психической информации (идеи, понятия, образы), носителями которых являются, в частности, нейронные ансамбли и их возбуждения в организме субъекта.

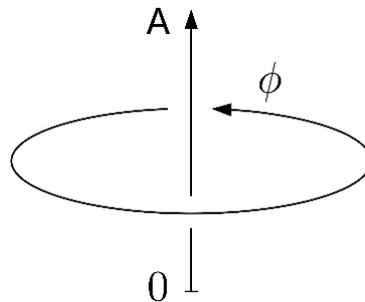


Рис. 6. Функциональное соотношение параметров когнитивных волн.

Амплитуда (вертикальная ось) кодирует объективно измеримую активность когнитивной единицы (идеи, понятия, образа), тогда как фаза (круговая координата) кодирует её субъективные интерференционно-смысловые отношения к другим когнитивным волнам

В информационной функции волн выделяются два разнокачественных аспекта, кодируемых амплитудой и фазой волны. Амплитуда волны соответствует размаху соответствующего нейронного колебания, кодируемую положительным действительным числом, как показано вертикальной осью на рис. 5. Эту величину можно получить, например, путём измерения картины электрического поля в головном мозге. При этом характерные амплитуды и конфигурации нейронных колебаний устойчивы настолько, что позволяют распознавать воспринимаемые человеком зрительные образы на основе электроэнцефалографических данных [61].

Фаза волны, напротив, для прямого измерения недоступна. Как видно из представленной модели, фазы ϕ_k кодируют не свойства когнитивных волн (б) самих по себе, а закономерности их композиции между собой. Эти параметры могут варьироваться независимо от объективных свойств соответствующих когнитивных единиц (например, образов самолёта или собаки, понятий о выигрыше или проигрыше, представляемых различными людьми более-менее единообразно). В этой связи можно предположить, что фазы когнитивных волн кодируют субъективный аспект психической информации, то есть *лично-смысловые* соотношения между образами, понятиями, идеями и другими психическими единицами. В силу независимости фазы волны от амплитуды, кольцевая топология фазовых параметров геометрически ортогональна амплитудной оси как показано на рис. 5. Полученная трёхмерная структура сходна с геометрией квантовой модели субъективного смысла [62].

Заключение. Представленные результаты показывают, что волновая логика является работоспособной альтернативой логике множеств, потенциально включающей последнюю в качестве предельного случая. В этой связи возникает вопрос о существовании других, возможно ещё более эффективных логик. Исключать такую возможность в принципе нельзя; однако, по принципу «что вверху – то и внизу» эти новые логики должны проявляться не только в мышлении человека, но и в простейших закономерностях природы. Физике же пока известен лишь один фундаментальный дуализм, образованный парами волна – частица и непрерывность – дискретность. В этой связи представляется, что дополнения логики множеств логикой волн может быть достаточно для значительного прогресса в моделировании субъективно-смысловых аспектов естественного поведения и мышления.

Список источников

1. Челпанов Г.И. Учебник Логики (1915)/ Г.И. Челпанов. – М.: Научная библиотека, 2010. – 128 с.
2. Колмогоров А.Н. Основные понятия теории вероятностей / А.Н. Колмогоров. – М. – Л. : ОНТИ, 1936. – 80 с.
3. von Neumann J., Morgenstern O. Theory of games and economic behavior. Princeton University Press, 1953.
4. Bouchaud J.P. Economics needs a scientific revolution. Nature, 2008, vol. 455, iss. 7217, pp. 1181.
5. Bouchaud J.P. The (unfortunate) complexity of the economy. Physics World, 2009, vol. 22, iss. 04, pp. 28-32.
6. Orrell D. Economyths: ten ways that economics gets it wrong. New Jersey, Wiley, 2010.
7. Stewart P. Why science does not get you. The nonergodic social world and the limit to measurement. World futures, 2022.
8. Khrennikov A. Ubiquitous quantum structure. From psychology to finance. Heidelberg, Springer, 2010, 216 p.
9. Суров И.А. Модели принятия решений в квантовой когнитивистике / И.А. Суров, А.П. Алоджанц. – Санкт-Петербург: Университет ИТМО, 2018. – 63 с.
10. Choustova O. Quantum probability and financial market. Information sciences, 2009, vol. 179, iss. 5.
11. Gonçalves C.P., Gonçalves C. An evolutionary quantum game model of financial market dynamics - theory and evidence. Capital markets: asset pricing and valuation (SSRN eJournal), 2008, vol. 31. iss. 11.
12. Khrennikov A. Towards information lasers. Entropy, 2015, vol. 17, iss. 10, pp. 6969-6994.
13. Khrennikov A. Social laser model: from color revolutions to Brexit and election of Donald Trump. Kybernetes, 2018, vol. 47, iss. 2, pp. 273-288.
14. Khrennikov A. Order stability via Fröhlich condensation in bio, eco, and social systems: The quantum-like approach. BioSystems, 2022, vol. 212, pp. 104593.
15. Khrennikov A. Open quantum systems in biology, cognitive and social sciences. Cham: Springer, 2023, 371 p.
16. Bradley R.T., Pribram K.H. Communication and stability in social collectives. Journal of Social and evolutionary systems, 1998, vol. 21, iss. 1, pp. 29-81.
17. Bradley R.T. Detecting the identity signature of secret social groups: Holographic processes and the communication of member affiliation. World futures: journal of general evolution, 2010, vol. 66, iss. 2, pp. 124-162.
18. Sallach D.L. Modeling emotional dynamics: currency versus field. Rationality and society, 2008, vol. 20, iss. 3, pp. 343-365.
19. Plikynas D. A virtual field-based conceptual framework for the simulation of complex social systems. Journal of Systems science and complexity, 2010, vol. 23, iss. 2, pp. 232-248.
20. Plikynas D., Basinskas G., Laukaitis A. Towards oscillations-based simulation of social systems: a neurodynamic approach. Connection science, 2015, vol. 27, iss. 2, pp. 188-211.
21. Orrell D. A Quantum theory of money and value. Economic thought, 2016, vol. 5, iss. 2, pp. 19-28.
22. Orrell D. The color of money: threshold effects in quantum economics. Quantum reports, 2021, vol. 3, iss. 2.
23. Orrell D. The value of value: A quantum approach to economics, security and international relations. Security dialogue, 2020, vol. 51, iss. 5, pp. 482-498.
24. Wendt A. The mind-body problem and social science: Motivating a quantum social theory. Journal for the Theory of social behaviour, 2018, vol. 48, iss. 2, pp. 188-204.
25. Shafir E., Tversky A. Thinking through uncertainty: Nonconsequential reasoning and choice. Cognitive psychology, 1992, vol. 24, iss. 4, pp. 449-474.
26. Tversky A., Shafir E. The disjunction effect in choice under uncertainty. Psychological science, 1992, vol. 3, iss. 5, pp. 305-309.
27. Kuhberger A., Komunská D., Perner J. The disjunction effect: does it exist for two-step gambles? Organizational behavior and human decision processes, 2001, vol. 85, iss. 2, pp. 250-264.
28. Lambdin C., Burdsal C. The disjunction effect reexamined: Relevant methodological issues and the fallacy of unspecified percentage comparisons. Organizational behavior and human decision processes, 2007, vol. 103, iss. 2, pp. 268-276.
29. Surov I.A., Pilkevich S.V., Alodjants A.P., Khmelevsky S.V. Quantum phase stability in human cognition. Frontiers in psychology, 2019, vol. 10.
30. Broekaert J.B., Busemeyer J.R., Pothos E.M. The disjunction effect in two-stage simulated gambles. An experimental study and comparison of a heuristic logistic, Markov and quantum-like model. Cognitive psychology, 2020, vol. 117, pp. 101262.
31. Croson R.T.A. The Disjunction effect and reason-based choice in games. Organizational behavior and human decision processes, 1999, vol. 80, iss. 2, pp. 118-133.
32. Busemeyer J.R., Matthews M., Wang Z. A quantum information processing explanation of disjunction effects. 28th Annual conference of the cognitive science society and the 5th International conference of cognitive science, 2006, iss. 1999, pp. 131-135.

33. Hristova E., Grinberg M. Disjunction effect in prisoner's dilemma: Evidences from an eye-tracking study. 30th Annual conference of the cognitive science society, 2008, pp. 1225-1230.
34. Li S., Taplin J.E. Examining whether there is a disjunction effect in Prisoner's Dilemma games. Chinese journal of Psychology, 2002, vol. 44, iss. 1, pp. 25-46.
35. Pothos E.M., Perry G., Corr P.J. Understanding cooperation in the Prisoner's Dilemma game. Personality and individual differences, 2011, vol. 51, iss. 3, pp. 210-215.
36. Khrennikov A. Quantum-like modeling of cognition. Frontiers in physics, 2015, vol. 3, iss. 77, pp. 77.
37. Zohuri B., Moghaddam M. What is fuzzy logic and how it works. In business resilience system (BRS): Driven Through boolean, fuzzy logics and cloud computation. Cham: Springer, 2017, pp. 199-219.
38. Kahraman C., Öztayşi B., Çevik Onar S. A comprehensive literature review of 50 years of fuzzy set theory. International journal of Computational intelligence systems, 2016, vol. 9, pp. 3-24.
39. Tversky A., Kahneman D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological review, 1983, vol. 90, iss. 4, pp. 293-315.
40. Kahneman D., Tversky A. Prospect theory. An analysis of decision under risk. Econometrica, 1979, vol. 47, iss. 2, pp. 263-291.
41. Erev I., Ert E., Plonsky O. From anomalies to forecasts: Toward a descriptive model of decisions under risk, under ambiguity, and from experience. Psychological review, 2017, vol. 124, iss. 4, pp. 369-409.
42. Blass A., Gurevich Y. Why sets? Pillars of computer science, Berlin, Springer, 2008, vol. 4800 LNCS, pp. 179-198.
43. Potter M. Set theory and its philosophy: a critical introduction, Oxford university Press, 2004.
44. Kauffman S.A., Gare A. Beyond descartes and newton: recovering life and humanity. Progress in biophysics and molecular biology, 2015, vol. 119, iss. 3, pp. 219-244.
45. Kauffman S., Roli A. The world is not a theorem. Entropy, 2021, vol. 23, iss. 11, pp. 1-9.
46. Иродов И.Е. Волновые процессы. Основные законы/ И.Е. Иродов. – М : Бинном, 2015. – 265 с.
47. Khrennikov A. Quantum-like cognition and rationality: biological and artificial intelligence systems. Quantum computing in the arts and humanities, Cham: Springer, 2022, pp. 153-178.
48. Yukalov V.I., Sornette D. Manipulating decision making of typical agents. IEEE Transactions on systems, man, and cybernetics: systems, 2014, vol. 44, iss. 9, pp. 1155-1168.
49. Shan Z. Brainwave phase stability: predictive modeling of irrational decision. Frontiers in psychology, 2022.
50. Barrett T.W. The cortex as interferometer: the transmission of amplitude, frequency and phase in cortical structures. Neuropsychologia, 1969, vol. 7, iss. 2, pp. 135-148.
51. Fries P. Rhythms for cognition: communication through coherence. Neuron, 2015, vol. 88, iss. 1, pp. 220–235.
52. Tiesinga P.H., Sejnowski T.J. Mechanisms for phase shifting in cortical networks and their role in communication through coherence. Frontiers in human neuroscience, 2010, vol. 4.
53. Neuling T. [et al.] Good vibrations: oscillatory phase shapes perception. NeuroImage, 2012, vol. 63, iss. 2, pp. 771–778.
54. Ten Oever S. [et al.] Phase-coded oscillatory ordering promotes the separation of closely matched representations to optimize perceptual discrimination. iScience, 2020, vol. 23, iss. 7.
55. Korotkov A.G., Levanova T.A., Zaks M.A., Maksimov A.G., Osipov G.V. Dynamics in a phase model of half-center oscillator: Two neurons with excitatory coupling. Communications in nonlinear science and numerical simulation, 2022, vol. 104.
56. Orefice A., Giovanelli R., Ditto D. The dynamics of wave-particle duality. Journal of Applied mathematics and physics, 2018, vol. 06, iss. 9, pp. 1840-1859.
57. Smith W.F. Waves and oscillations a prelude to quantum mechanics. Oxford University Press, 2010.
58. Мандель Л. Оптическая когерентность и квантовая оптика / Мандель Л., Вольф Э. – М.: Физматлит, 2000. – 896 с.
59. Мелик-Гайказян И.В. Информационные процессы и реальность / И.В. Мелик-Гайказян. – М : Наука, 1998. – 192 с.
60. Корогодина В.И. Информация как основа жизни / В.И. Корогодина, В.Л. Корогодина. – Дубна : Издательский центр «Феникс», 2000. – 208 с.
61. Kavasidis I., Palazzo S., Spampinato pp., Giordano D., Shah M. Brain2Image: converting brain signals into images. Proceedings of the 25th ACM international conference on multimedia, ACM, 2017, pp. 1809-1817.
62. Суров И.А. Какая разница? Прагматическая формализация смысла / И.А. Суров // Искусственный интеллект и принятие решений, 2023. – № 1. – С. 78-89.

Суров Илья Алексеевич. Кандидат физ.-мат. наук, доцент, университет ИТМО. Область научных интересов: когнитивно-поведенческое моделирование, квантовая семантика. AuthorID: 819354, SPIN: 5003-7439, ORCID: 0000-0001-5690-7507, ilya.a.surov@itmo.ru, 197101, Россия, Санкт-Петербург, Кронверкский проспект, 49А.

Logic of sets and logic of waves in cognitive-behavioral modeling**Ilya A. Surov**ITMO University, Russia, Saint Petersburg, *ilya.a.surov@itmo.ru*

Abstract. Cognitive-behavioral processes routinely deviate from classical models, assuming the deterministic rationality of actors. The paper revises the logical foundations behind these limitations. A major such premise is shown to be the use of classical probability theory and the logic of sets (including fuzzy ones) behind it, sometimes mistakenly considered as the only option. According to the natural dualism of discreteness – continuity and particle – wave, another possibility is the logic of waves and the corresponding probability calculus. The efficiency of this logic in cognitive-behavioral modeling is demonstrated in the experiments «prisoner’s dilemma» and «two-stage gamble». A wave-like probabilistic model complements the logic of sets with an additional factor, quantifying violations of classical rationality in these experiments. Phases of the interfering cognitive waves in this model account for the subjectively-semantic regularities of natural thinking, ignored in classical approaches. These new regularities enable probabilistic forecast of «irrational» decisions in novel contexts. Unique features of wave logic, common to optical, holographic, and quantum algorithms of information, open novel prospects for cognitive-behavioral modeling and data analysis.

Keywords: wave logic, cognitive interference, behavioral modeling, probabilistic prognosis, set calculus, rationality

References

1. Chelpanov G.I. Uchebnyk Logiki [Tutorial in Logic] (1915). Moscow: Nauchnaya biblioteka [Science Library], 2010. 128 p.
2. Kolmogorov A.N. Osnovnyye ponyatiya teorii veroyatnostey [Basic concepts of probability theory]. M., L., ONTI, 1936, 80 p.
3. von Neumann J., Morgenstern O. Theory of games and economic behavior. Princeton University Press, 1953.
4. Bouchaud J.P. Economics needs a scientific revolution. *Nature*, 2008, vol. 455, iss. 7217, pp. 1181.
5. Bouchaud J.P. The (unfortunate) complexity of the economy. *Physics World*, 2009, vol. 22, iss. 04, pp. 28-32.
6. Orrell D. Economyths: ten ways that economics gets it wrong. New Jersey, Wiley, 2010.
7. Stewart P. Why science does not get you. The nonergodic social world and the limit to measurement. *World futures*, 2022.
8. Khrennikov A. Ubiquitous quantum structure. From psychology to finance. Heidelberg, Springer, 2010, 216 p.
9. Surov I.A., Alodjants A.P. Modeli prinyatiya reshenij v kvantovoj kognitivistike [Models of decision making in quantum cognition]. Saint-Petersburg, ITMO University, 2018, 63 p.
10. Choustova O. Quantum probability and financial market. *Information sciences*, 2009, vol. 179, iss. 5.
11. Gonçalves C.P., Gonçalves C. An evolutionary quantum game model of financial market dynamics - theory and evidence. *Capital markets: asset pricing and valuation (SSRN eJournal)*, 2008, vol. 31, iss. 11.
12. Khrennikov A. Towards information lasers. *Entropy*, 2015, vol. 17, iss. 10, pp. 6969-6994.
13. Khrennikov A. Social laser model: from color revolutions to Brexit and election of Donald Trump. *Kybernetes*, 2018, vol. 47, iss. 2, pp. 273-288.
14. Khrennikov A. Order stability via Fröhlich condensation in bio, eco, and social systems: The quantum-like approach. *BioSystems*, 2022, vol. 212, pp. 104593.
15. Khrennikov A. Open quantum systems in biology, cognitive and social sciences. Cham: Springer, 2023. 371 p.
16. Bradley R.T., Pribram K.H. Communication and stability in social collectives. *Journal of Social and evolutionary systems*, 1998, vol. 21, iss. 1, pp. 29-81.
17. Bradley R.T. Detecting the identity signature of secret social groups: Holographic processes and the communication of member affiliation. *World futures: journal of general evolution*, 2010, vol. 66, iss. 2, pp. 124-162.
18. Sallach D.L. Modeling emotional dynamics: currency versus field. *Rationality and society*, 2008, vol. 20, iss. 3, pp. 343-365.
19. Plikynas D. A virtual field-based conceptual framework for the simulation of complex social systems. *Journal of Systems science and complexity*, 2010, vol. 23, iss. 2, pp. 232-248.
20. Plikynas D., Basinskas G., Laukaitis A. Towards oscillations-based simulation of social systems: a neurodynamic approach. *Connection science*, 2015, vol. 27, iss. 2, pp. 188-211.
21. Orrell D. A Quantum theory of money and value. *Economic thought*, 2016, vol. 5, iss. 2, pp. 19-28.
22. Orrell D. The color of money: threshold effects in quantum economics. *Quantum reports*, 2021, vol. 3, iss. 2.

23. Orrell D. The value of value: A quantum approach to economics, security and international relations. *Security dialogue*, 2020, vol. 51, iss. 5, pp. 482-498.
24. Wendt A. The mind–body problem and social science: Motivating a quantum social theory. *Journal for the Theory of social behaviour*, 2018, vol. 48, iss. 2, pp. 188-204.
25. Shafir E., Tversky A. Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive psychology*, 1992, vol. 24, iss. 4, pp. 449-474.
26. Tversky A., Shafir E. The disjunction effect in choice under uncertainty. *Psychological science*, 1992, vol. 3, iss. 5, pp. 305-309.
27. Kuhberger A., Komunská D., Perner J. The disjunction effect: does it exist for two-step gambles? *Organizational behavior and human decision processes*, 2001, vol. 85, iss. 2, pp. 250-264.
28. Lambdin C., Burdsal C. The disjunction effect reexamined: Relevant methodological issues and the fallacy of unspecified percentage comparisons. *Organizational behavior and human decision processes*, 2007, vol. 103, iss. 2, pp. 268-276.
29. Surov I.A., Pilkevich S.V., Alodjants A.P., Khmelevsky S.V. Quantum phase stability in human cognition. *Frontiers in psychology*, 2019, vol. 10.
30. Broekaert J.B., Busemeyer J.R., Pothos E.M. The disjunction effect in two-stage simulated gambles. An experimental study and comparison of a heuristic logistic, Markov and quantum-like model. *Cognitive psychology*, 2020, vol. 117, pp. 101262.
31. Croson R.T.A. The Disjunction effect and reason-based choice in games. *Organizational behavior and human decision processes*, 1999, vol. 80, iss. 2, pp. 118-133.
32. Busemeyer J.R., Matthews M., Wang Z. A quantum information processing explanation of disjunction effects. 28th Annual conference of the cognitive science society and the 5th International conference of cognitive science, 2006, iss. 1999, pp. 131-135.
33. Hristova E., Grinberg M. Disjunction effect in prisoner's dilemma: Evidences from an eye-tracking study. 30th Annual conference of the cognitive science society, 2008, pp. 1225-1230.
34. Li S., Taplin J.E. Examining whether there is a disjunction effect in Prisoner's Dilemma games. *Chinese journal of Psychology*, 2002, vol. 44, iss. 1, pp. 25-46.
35. Pothos E.M., Perry G., Corr P.J. Understanding cooperation in the Prisoner's Dilemma game. *Personality and individual differences*, 2011, vol. 51, iss. 3, pp. 210-215.
36. Khrennikov A. Quantum-like modeling of cognition. *Frontiers in physics*, 2015, vol. 3, iss. 77, pp. 77.
37. Zohuri B., Moghaddam M. What is fuzzy logic and how it works. In *business resilience system (BRS): Driven Through boolean, fuzzy logics and cloud computation*. Cham: Springer, 2017, pp. 199-219.
38. Kahraman C., Öztayşi B., Çevik Onar S. A comprehensive literature review of 50 years of fuzzy set theory. *International journal of Computational intelligence systems*, 2016, vol. 9, pp. 3-24.
39. Tversky A., Kahneman D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review*, 1983, vol. 90, iss. 4, pp. 293-315.
40. Kahneman D., Tversky A. Prospect theory. An analysis of decision under risk. *Econometrica*, 1979, vol. 47, iss. 2, pp. 263-291.
41. Erev I., Ert E., Plonsky O. From anomalies to forecasts: Toward a descriptive model of decisions under risk, under ambiguity, and from experience. *Psychological review*, 2017, vol. 124, iss. 4, pp. 369-409.
42. Blass A., Gurevich Y. *Why sets? Pillars of computer science*, Berlin, Springer, 2008, vol. 4800 LNCS, pp. 179-198.
43. Potter M. *Set theory and its philosophy: a critical introduction*, Oxford university Press, 2004.
44. Kauffman S.A., Gare A. Beyond descartes and newton: recovering life and humanity. *Progress in biophysics and molecular biology*, 2015, vol. 119, iss. 3, pp. 219-244.
45. Kauffman S., Roli A. The world is not a theorem. *Entropy*, 2021, vol. 23, iss. 11, pp. 1-9.
46. Irodov I.E. *Volnovye processy. Osnovnye zakony [Wave processes. Basic laws]*. M., Binom, 2015, 265 p.
47. Khrennikov A. Quantum-like cognition and rationality: biological and artificial intelligence systems. *Quantum computing in the arts and humanities*, Cham: Springer, 2022, pp. 153-178.
48. Yukalov V.I., Sornette D. Manipulating decision making of typical agents. *IEEE Transactions on systems, man, and cybernetics: systems*, 2014, vol. 44, iss. 9, pp. 1155-1168.
49. Shan Z. Brainwave phase stability: predictive modeling of irrational decision. *Frontiers in psychology*, 2022.
50. Barrett T.W. The cortex as interferometer: the transmission of amplitude, frequency and phase in cortical structures. *Neuropsychologia*, 1969, vol. 7, iss. 2, pp. 135-148.
51. Fries P. Rhythms for cognition: communication through coherence. *Neuron*, 2015, vol. 88, iss. 1, pp. 220–235.
52. Tiesinga P.H., Sejnowski T.J. Mechanisms for phase shifting in cortical networks and their role in communication through coherence. *Frontiers in human neuroscience*, 2010, vol. 4.

53. Neuling T. [et al.] Good vibrations: oscillatory phase shapes perception. *NeuroImage*, 2012, vol. 63, iss. 2, pp. 771–778.
54. Ten Oever S. [et al.] Phase-coded oscillatory ordering promotes the separation of closely matched representations to optimize perceptual discrimination. *iScience*, 2020, vol. 23, iss. 7.
55. Korotkov A.G., Levanova T.A., Zaks M.A., Maksimov A.G., Osipov G.V. Dynamics in a phase model of half-center oscillator: Two neurons with excitatory coupling. *Communications in nonlinear science and numerical simulation*, 2022, vol. 104.
56. Orefice A., Giovanelli R., Ditto D. The dynamics of wave-particle duality. *Journal of Applied mathematics and physics*, 2018, vol. 06, iss. 9, pp. 1840-1859.
57. Smith W.F. *Waves and oscillations a prelude to quantum mechanics*. Oxford University Press, 2010
58. Mandel L., Wolf E. *Opticheskaya kogerentnost' i kvantovaya optika [Optical coherence and quantum optics]*. М., Fizmatlit, 2000, 896 p.
59. Melik-Gajkazyan I.V. *Informacionnye processy i real'nost' [Information processes and reality]*. Moscow, Nauka [Science], 1998, 192 p.
60. Korogodin V.I., Korogodina V.L. *Informaciya kak osnova zhizni [Information as the basis of life]*. Dubna, Izdatel'skiy tsentr "Feniks" [Publishing center "Phoenix"], 2000, 208 p.
61. Kavasidis I., Palazzo S., Spampinato pp., Giordano D., Shah M. Brain2Image: converting brain signals into images. *Proceedings of the 25th ACM international conference on multimedia*, ACM, 2017, pp. 1809-1817.
62. Surov I.A. Kakaya raznica? Pragmaticheskaya formalizaciya smysla [What is the difference? Pragmatic formalization of meaning]. *Iskusstvennyy intellekt i prinyatiye resheniy [Artificial intelligence and decision making]*, 2023, iss. 1, pp. 78-89.

Surov Ilya Alekseevich. *PhD in physics and mathematics, associate professor, ITMO university. Research interests: cognitive modeling, quantum semantics. AuthorID: 819354, SPIN: 5003-7439, ORCID: 0000-0001-5690-7507, ilya.a.surov@itmo.ru, 197101 Russia, Saint Petersburg, Kronverksky Prospekt, 49A.*

Статья поступила в редакцию 11.07.2023; одобрена после рецензирования 08.12.2023; принята к публикации 16.12.2023.

The article was submitted 07/11/2023; approved after reviewing 12/08/2023; accepted for publication 12/16/2023.

УДК 004.89+004.048

DOI: 10.25729/ESI.2023.32.4.006

Применение машинного обучения для анализа образовательных результатов студентов вузов

Алпатов Алексей Викторович

Волжский политехнический институт (филиал) Волгоградского государственного технического университета, Волжский, Россия, *alpatov80@mail.ru*

Аннотация. В работе представлены результаты анализа и прогнозирования образовательных результатов студентов первого курса вуза при реализации отдельной дисциплины с использованием машинного обучения. Актуальность темы исследования обусловлена необходимостью вузов в современных условиях успешно конкурировать на рынке образовательных услуг, который характеризуется низким количеством абитуриентов и увеличением требований к качеству профессионального образования, как со стороны абитуриентов, так и со стороны государства. Важной составляющей для эффективного принятия решения в процессе управления качеством образовательного процесса является учебная аналитика, на основе которой можно выполнять прогнозирование академической успеваемости студентов, выявлять факторы, оказывающие существенное влияние на достижение высоких образовательных результатов. Исследование показало возможность прогнозирования сдачи экзамена по отдельной дисциплине студентов первого курса вуза на основе данных контрольных срезов, которые проводят деканаты в течение семестра для выявления групп студентов с повышенным риском возникновения академической задолженности. Точность прогнозирования, которую показали построенные модели (наивный байесовский классификатор и логистическая регрессия) оказалась вполне приемлемой, как на этапе проведения первого рубежного контроля, так и на этапе проведения второго. Результаты данной работы имеют практическое значение для администрации вузов и для преподавателей. Прогнозные модели можно использовать при прогнозировании отчисления студентов вследствие академической неуспеваемости. Модели могут быть встроены в образовательные информационные системы и быть помощником преподавателям для принятия решений в процессе реализации дисциплины.

Ключевые слова: прогнозирование образовательных результатов, учебная аналитика, интеллектуальный анализ образовательных данных, единый государственный экзамен, нейронные сети, дерево решений, логистическая регрессия, кластеризация

Цитирование: Алпатов А.В. Применение машинного обучения для анализа образовательных результатов студентов вузов / А.В. Алпатов // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 67-78. – DOI: 10.25729/ESI.2023.32.4.006.

Введение. В сфере профессионального образования одной из приоритетных задач является поддержание высокого уровня образования с сохранением достаточно большой доли выпускников среди поступивших в учебное заведение. Это особенно актуально в современных условиях российского рынка образования, который характеризуется низким количеством абитуриентов и увеличением требований к качеству профессионального образования, как со стороны абитуриентов, так и со стороны государства. Важной составляющей для эффективного принятия решения в процессе управления качеством образовательного процесса является учебная аналитика, на основе которой можно выполнять прогнозирование академической успеваемости студентов, выявлять факторы, оказывающих существенное влияние на достижение высоких образовательных результатов.

Анализ научных публикаций, посвященных интеллектуальному анализу академических успехов студентов в высших учебных заведениях, показал, что существенное внимание уделено вопросам прогнозирования отчисления студентов [1, 2], выявления группы риска студентов образования академической задолженности по отдельной дисциплине [3, 4, 5], а также прогнозирования величины среднего балла [6]. Факторы, которые исследователи выбирают для построения прогнозных моделей академической успеваемости, можно разделить на такие группы, как образовательные, материальные, социальные, интеллектуальные и мотивационные. Ряд исследований посвящен изучению влияния различных факторов на образовательные

результаты. Например, воздействие социальных и демографических характеристик рассматривается в статье [7], в публикациях [8, 9] представлены результаты исследования по влиянию мотивации студентов на их образовательные результаты.

С развитием концепции индивидуальных образовательных траекторий, а также широким внедрением электронных образовательных платформ одним из направлений научных исследований стало прогнозирование образовательных траекторий отдельных студентов [4, 10, 11, 12]. В частности, в работе [11] рассматривался вопрос построения моделей прогнозирования успешности по дисциплине, которые позволяют в динамике «выявлять студентов с повышенным риском не аттестации по дисциплине». Прогнозирование осуществлялось на основе ежедневных данных об успеваемости студентов и активности в электронной образовательной среде с использованием различных методов машинного обучения. Наилучшую точность при прогнозировании показала модель Ансамбль – усреднение классификаторов Random Forest, XGBoost и Logistic Regression с регуляризацией L1. При этом в работе отмечено, что «данные цифрового следа уже в первой половине семестра позволяют выявлять студентов с высоким риском не успешности по дисциплине».

В настоящее время некоторые российские вузы вводят информационные образовательные системы для прогнозирования образовательных результатов [11]. Однако у большинства учебных заведений обычной практикой при анализе успеваемости студентов в течение семестра является проведение одного или двух контрольных срезов (рубежных контролей). Как правило, в этих срезах указываются данные о набранных баллах и о количестве пропусков. Эти сведения по всем дисциплинам передаются в деканат, объединяются и на основе агрегатных статистических показателей деканат выявляет студентов, которые находятся в группе риска на образование академической задолженности.

Целью данной работы является исследование возможности использования данных, полученных на контрольных срезах, для анализа и прогнозирования образовательных результатов студентов первого курса с помощью методов машинного обучения. Прогнозирование академических успехов обучающихся представляет собой задачу бинарной классификации. В качестве предиктивной переменной выступает дамми-переменная, которая принимает значение 1, если студент сдал экзамен на любую положительную оценку и 0, если студент не сдал экзамен. Еще одной задачей в данной работе является выявление на основе кластерного анализа групп студентов со схожими итоговыми образовательными результатами по отдельной дисциплине.

1. Данные и методы. Для построения прогностических моделей были использованы данные, которые представляли собой результаты контрольных срезов по дисциплине «Основы программирования» за первый семестр студентов очной формы обучения направлений подготовки 09.03.01 Информатика и вычислительная техника, а также 09.03.09 Программная инженерия. Моделирование осуществлялось на данных о 129 студентах, которые поступили в Волжский политехнический институт в 2020, 2021 и 2022 году.

Для обучения моделей использовались также результаты вступительных испытаний студентов при поступлении в Институт. В качестве вступительного испытания может быть принят результат Единого государственного экзамена (ЕГЭ), если студент поступил на базе среднего общего образования или результат экзамена, проводимого Институтом, если студент поступил на базе среднего профессионального образования. В большинстве исследований, в которых рассматривается вопрос прогнозирования образовательных результатов российских студентов, основным фактором влияния на успешность обучения являются результаты ЕГЭ [6; 13]. Однако, из ряда исследований, например, [14], известно, что результаты ЕГЭ наиболее сильное влияние оказывают на первых курсах обучения, а в дальнейшем эта связь уменьшается. В данной работе анализируются образовательные данные студентов первого курса, поэтому результаты вступительных испытаний предположительно будут оказывать наибольшее

влияние на успешность освоения учебных дисциплин, и их можно рассматривать как своего рода потенциал обучающегося.

Для мониторинга ритмичной работы студентов деканат в Институте проводит в течение семестра два контрольных среза (рубежных контроля) на шестой и двенадцатой неделях графика обучения. Данные о набранных баллах и количестве пропущенных студентами аудиторных часов, которые преподаватели передают в деканат, представлены в виде относительных величин. В таблице 1 представлены обозначения переменных, используемых при моделировании образовательных результатов, и их описание.

Таблица 1. Описание и обозначения переменных

Обозначение	Название переменной	Описание переменной
<i>Exam</i>	Результат сдачи экзамена по дисциплине	Бинарная переменная, которая принимает значение 1, если студент сдал экзамен и 0, если студент его не сдал
<i>Group</i>	Направление подготовки	Бинарная переменная, которая принимает значение 1, если студент обучается по направлению подготовки 09.03.09 Программная инженерия и 0, если студент обучается по направлению подготовки 09.03.01 Информатика и вычислительная техника
<i>Sex</i>	Пол студента	Бинарная переменная, которая принимает значение 1, если пол мужской и 0, если пол женский
<i>Type_exam</i>	Вид вступительных испытаний в институт	Бинарная переменная, которая принимает значение 1, если абитуриент был зачислен в институт по результатам ЕГЭ и 0, если он поступил в институт на основе экзамена, проводимого вузом
<i>Entrance_exam</i>	Результаты вступительного экзамена	Переменная <i>Entrance_exam</i> представляет собой отношение суммы баллов за три вступительных испытания к их максимально возможной сумме, т.е. к 300
<i>Academ1</i>	Успеваемость к первому рубежному контролю	Доля баллов, набранных на этапе первого рубежного контроля, от максимально возможного количества баллов к данному этапу
<i>Academ2</i>	Успеваемость ко второму рубежному контролю	Доля баллов, набранных на этапе второго рубежного контроля, от максимально возможного количества баллов к данному этапу
<i>Attend1</i>	Посещаемость к первому рубежному контролю	Отношение количества посещений учебных занятий студентом к числу аудиторных занятий, проведенных на момент первого рубежного контроля
<i>Attend2</i>	Посещаемость ко второму рубежному контролю	Отношение количества посещений учебных занятий студентом к числу аудиторных занятий, проведенных на момент второго рубежного контроля
<i>AcademF</i>	Итоговая успеваемость за первый семестр	Доля баллов, набранных в течение первого семестра от максимально возможного
<i>AttendF</i>	Итоговая посещаемость за первый семестр	Отношение количества посещений учебных занятий студентом за первый семестр к числу аудиторных занятий, проведенных в течение семестра

Прогнозирование осуществляется дважды: на момент проведения первого и второго рубежного контроля. Показатели, характеризующие успеваемость и посещаемость, аккумулируют все другие возможные факторы, которые формируют знания по дисциплине. В связи с этим нет необходимости проведения дополнительного тестирования, сбора данных, которые касаются личности студента и его деятельности в процессе обучения.

При построении прогностических моделей широко используются методы машинного обучения. Как правило, исследователи строят несколько моделей и сравнивают точность прогнозирования. Например, в работе [15] применялись модели дерева решений (Decision Tree), случайный лес (Random Forest), наивный Байес (Naive Bayes), а также индукция правил (Rule induction). В целом наилучший результат показал метод дерева решений. В рамках проводимого исследования прогнозирование в задаче классификации осуществлялось на основе двух моделей: наивный байесовский классификатор и логистическая регрессия. Выбор был обусловлен тем, что объем выборки небольшой и наилучшие результаты при обучении будут давать простые модели. Для выявления закономерностей при формировании результатов обучения за первый семестр был проведен кластерный анализ с использованием иерархического подхода. При моделировании использовались библиотеки Python: Keras, sklearn, scipy. При проверке статистических гипотез применялись тесты из модуля stats библиотеки scipy.

2. Разведочный анализ данных. Проведем статистический анализ набора данных, который будет использован при прогнозировании. Анализируемый набор данных содержит 52,7% обучающихся по направлению подготовки «Программная инженерия» и, соответственно, 47,3% по направлению подготовки «Информатика и вычислительная техника». При этом, доля студенток составляет 13,2%, доля поступивших на базе среднего профессионального образования 7,0%. По итогу освоения дисциплины в первом семестре в среднем экзамен сдают 59,7% студентов обоих направлений подготовки. Таким образом, в классах присутствует дисбаланс, но он незначительный.

В таблице 2 приведены основные описательные статистики для количественных показателей. Значения статистических показателей переменной *Entrance_exam* представлены в процентах и в баллах. Средний балл вступительных испытаний при поступлении на специальности «Программная инженерия» и «Информатика и вычислительная техника» составляет 193,5 со среднеквадратическим отклонением 24,0. Крайние значения изменяются от 127,0 до 264,0 баллов. На рис. 1 показана гистограмма распределения результатов вступительных испытаний студентов, представленных в датасете для двух групп студентов: сдавших экзамен и не сдавших его. Визуально распределение для сдавших экзамен близко к нормальному. Распределение для $Exam = 0$ ограничено значением $Entrance_exam = 0,74$. Это означает, что студенты, которые набрали примерно более 222 баллов, всегда сдают экзамен вовремя.

Таблица 2. Описательная статистика

Показатель	<i>Entrance_exam</i>		<i>Academ1</i>	<i>Attend1</i>	<i>Academ2</i>	<i>Attend2</i>
	доля	балл				
mean	0,645	193,5	0,522	0,826	0,519	0,758
std	0,080	24,0	0,242	0,268	0,254	0,279
min	0,423	127,0	0,000	0,000	0,000	0,000
25%	0,593	178,0	0,375	0,800	0,353	0,667
50%	0,643	193,0	0,538	0,933	0,529	0,857
75%	0,700	210,0	0,706	1,000	0,711	0,966
max	0,880	264,0	1,000	1,000	0,974	1,000

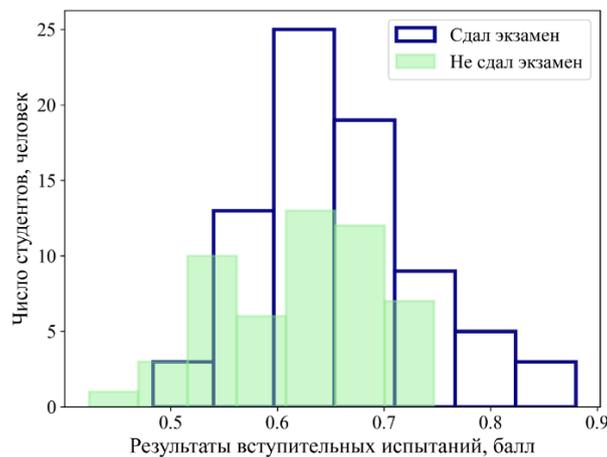


Рис. 1. Распределения результатов вступительных испытаний студентов, поступивших на направления подготовки «Программная инженерия» и «Информатика и вычислительная техника»

В таблице 3 показаны средние значения *mean* и среднеквадратические отклонения *std* переменной *Entrance_exam* для различных классов студентов. Можно отметить, что средний балл студенток, поступивших в вуз, выше, чем у студентов (203,1 и 192,1, соответственно). Проверка гипотезы о равенстве средних с помощью *t*-теста показала, что нулевая гипотеза отклоняется на уровне значимости 0,05 (*P*-значение = 0,039). При проверке данной гипотезы предполагалось, что обе выборки имеют распределения, близкие к нормальному. Поскольку количество студенток небольшое, то для проверки гипотезы о нормальности использовался критерий Шапиро. *P*-значение оказалось равным 0,064 и нулевая гипотеза не отклоняется, если уровень значимости принять равным 0,05. Для проверки гипотезы о равенстве средних использовался критерий Флигнера-Килина (*P*-значение = 0,743). На основе результатов вступительного экзамена можно сделать предположение о том, что базовые знания у девушек первого курса в целом выше, чем у юношей (однако данное предположение не является в достаточной степени надежным). Среднее значение баллов за вступительные испытания у студентов направления подготовки «Программная инженерия» существенно выше, чем у студентов, обучающихся по направлению «Информатика и вычислительная техника». Проведенный *t*-тест о равенстве средних подтвердил данное различие (*P*-значение < 0,001). Таким образом, уровень подготовки студентов, поступивших на направление подготовки «Программная инженерия», превышает уровень подготовки тех, кто поступил на направление «Информатика и вычислительная техника». Среднее значение *Entrance_exam* для сдавших экзамен по дисциплине, ожидаемо выше почти на 11 баллов по сравнению с теми, кто не сдал экзамен.

Таблица 3. Результаты вступительного испытания, балл

Показатель	<i>Sex</i>		<i>Group</i>		<i>Exam</i>	
	<i>1</i>	<i>0</i>	<i>1</i>	<i>0</i>	<i>1</i>	<i>0</i>
mean	192,1	203,1	199,9	186,5	197,9	187,0
std	23,7	24,3	21,8	24,4	24,2	22,3

В таблице 4 представлены значения коэффициентов корреляции между различными переменными анализируемого набора данных. Рассмотрим линейную корреляцию между прогнозируемым признаком *Exam* и факторными признаками. Наиболее сильное влияние оказывают показатели, характеризующие академическую успеваемость и посещаемость учебных занятий. При этом для второго рубежного контроля эта связь выражена в большей степени, вследствие близости промежуточной аттестации по дисциплине. Корреляции *Exam* с переменной *Sex* слабая и отрицательная, что может свидетельствовать о том, что студентки имеют не намного больше шансов сдать экзамен. Линейная связь между прогнозируемой переменной и

переменной *Type_exam* практически отсутствует. Проверка статистической значимости коэффициента корреляции показала, что нулевая гипотеза на уровне значимости 0,01 не отклоняется. В связи с этим данная переменная была исключена из датасета.

Таблица 4. Матрица коэффициентов корреляции

	<i>Group</i>	<i>Sex</i>	<i>Type_exam</i>	<i>Entrance_exam</i>	<i>Academ1</i>	<i>Attend1</i>	<i>Academ2</i>	<i>Attend2</i>	<i>Exam</i>
<i>Group</i>	1	-0,185	0,167	0,279	0,277	0,100	0,213	0,197	0,330
<i>Sex</i>	-0,185	1	-0,107	-0,155	-0,098	-0,123	-0,069	-0,157	-0,18
<i>Type_exam</i>	0,167	-0,107	1	0,042	0,054	0,036	-0,065	0,075	-0,039
<i>Entrance_exam</i>	0,279	-0,155	0,042	1	0,036	0,027	0,119	0,128	0,224
<i>Academ1</i>	0,277	-0,098	0,054	0,036	1	0,798	0,792	0,743	0,59
<i>Attend1</i>	0,100	-0,123	0,036	0,027	0,798	1	0,716	0,887	0,523
<i>Academ2</i>	0,213	-0,069	-0,065	0,119	0,792	0,716	1	0,847	0,679
<i>Attend2</i>	0,197	-0,157	0,075	0,128	0,743	0,887	0,847	1	0,672
<i>Exam</i>	0,330	-0,180	-0,039	0,224	0,59	0,523	0,679	0,672	1

Коэффициенты корреляции между посещаемостью и успеваемостью довольно высокие, что является вполне закономерным, поскольку высокий уровень посещаемости, как правило, свидетельствует о высокой мотивации студентов к получению знаний и навыков по дисциплине «Основы программирования». Кроме того, основная часть баллов, набирается при проведении аудиторных занятий. Включение коррелирующих факторов, например, *Academ1* и *Attend1*, в линейную регрессионную модель приведет к появлению мультиколлинеарности. Поскольку корреляция с прогнозируемой переменной *Exam* выше у тех переменных, которые показывают долю набранных баллов к рубежному контролю, то в прогнозные модели будут включаться только *Academ1* и *Academ2*, а переменные *Attend1* и *Attend2* не будут включены. Связь между прогнозируемой переменной и переменной *Entrance_exam* довольно слабая и составляет 0,224. Это обусловлено несколькими причинами. Во-первых, в качестве вступительных испытаний принимались результаты ЕГЭ или результаты внутреннего экзамена Института. Базовой дисциплиной ЕГЭ может быть одна из трех дисциплин: информатика, физика или английский язык. Поэтому по профильным дисциплинам у студентов базовый уровень подготовки может отличаться. Во-вторых, студенты первого курса имеют разный опыт программирования до поступления в вуз.

3. Прогнозные модели. При прогнозировании переменной *Exam* на момент проведения первого рубежного контроля в качестве факторных признаков были использованы переменные *Group*, *Sex*, *Entrance_exam*, *Academ1*. На этапе прохождения второго рубежного – *Group*, *Sex*, *Entrance_exam*, *Academ2*.

Для сравнения были получены разные модели: наивный байесовский метод и логистическая регрессия. При обучении моделей данные делились на обучающую и тестовую выборки. Доля данных в тестовой выборке составляла 20 %.

Рассмотрим вопрос выбора метрики для оценки качества полученных моделей. При прогнозировании бинарной переменной возможны ошибки двух типов: ошибки первого рода (False Positive) и ошибки второго рода (False Negative). В рассматриваемой задаче ошибка первого рода заключается в неверном предсказании того, что студент сдаст экзамен по дисциплине. Соответственно, ошибка второго рода состоит в неверном предсказании того, что студент не сдаст экзамен. В таблице 5 показана матрица ошибок для задачи прогнозирования сдачи экзамена по дисциплине.

Таблица 5. Матрица ошибок

		Фактическое значение FY	
		1	0
Прогнозное значение FY	1	True Positive (TP) Верный прогноз, что студент сдаст экзамен	False Positive (FP) Неверный прогноз, что студент сдаст экзамен
	0	False Negative (FN) Неверный прогноз, что студент не сдаст экзамен	True Negative (TN) Верный прогноз, что студент не сдаст экзамену

Для оценки качества работы алгоритмов в задачах бинарной классификации часто используются такие метрики, как меткость *Accuracy*, точность *Precision* и полнота *Recall*.

Precision можно интерпретировать, как долю студентов, для которых модель сделала верный прогноз о сдаче экзамена. Чем выше значение *Precision*, тем ниже доля ошибок первого рода. Метрика *Recall* показывает долю верно прогнозируемых сдач экзамена среди всех фактически сдавших экзамен студентов. Чем выше *Recall*, тем ниже доля ошибок второго рода. Метрика *Accuracy* актуальна в том случае, когда оба класса имеют одинаковое значение для исследователя. В рассматриваемой задаче наиболее важным является контролировать ошибки первого рода, поскольку в этом случае последствия для процесса управления численностью контингента будут более негативными. Ведь в этом случае заведомо неуспевающий студент будет признан успевающим и на него не будет оказано вовремя воспитательное воздействие. Безусловно, желательно не оставлять без внимания и ошибки второго рода, поскольку при наличии большого числа обучающихся преподаватель будет тратить существенное количество времени для мотивации студентов, которые являются успевающими. Лучше сконцентрировать больше внимания и усилий для работы с действительно отстающими.

С учетом данных аргументов для сравнения качества моделей была выбрана F -мера с β – коэффициентом равным 0,5:

$$F_{\beta} = (1 + \beta^2) \cdot \frac{Precision \cdot Recall}{(\beta^2 \cdot Precision) + Recall}$$

В таблице 6 представлены показатели качества моделей прогнозирования, *Precision* и F -мера для тестовой выборки. Лучший результат показала модель «дерево решений». Можно отметить, что прогнозирование уже на первом рубежном контроле имеет приемлемую точность. Это дает возможность на ранних этапах обучения выявлять отстающих студентов. При обучении на данных, полученных на втором рубежном контроле, модели оказались немного точнее, что объясняется близостью окончания семестра.

4. Кластерный анализ. Для выявления закономерностей в результатах обучения студентов за первый семестр был проведен иерархический кластерный анализ. Использовались характеристики объектов: *Academ1*, *Attend1*, *Academ2*, *Attend2*, *Sex*, *Exam*, *Group*, итоговая успеваемость за первый семестр *AcademF*, итоговая посещаемость за первый семестр *AttendF*, а также бинарная переменная Y .

Таблица 6. F -мера ($\beta = 0.5$) различных моделей прогнозирования на тренировочной и тестовой выборках

Модель	Первый рубежный контроль		Второй рубежный контроль	
	Тренировочная	Тестовая	Тренировочная	Тестовая
Наивный байесовский классификатор	0.83	0.82	0.85	0.89
Логистическая регрессия	0,82	0,84	0,85	0,88

В процессе выявления кластеров комбинировались различные метрики оценки расстояний между объектами и метрики расчета расстояний между кластерами. Анализировались различные варианты разбиения объектов на группы. Было принято решение остановиться на 6 кластерах, поскольку при таком делении результаты кластеризации можно проинтерпретировать наилучшим образом. Данное разбиение было получено при выборе метода Варда (ward) для оценки степени сходства объектов евклидового расстояния (euclidean) и для оценки расстояний между кластерами. На рис. 2 показана дендрограмма, которая иллюстрирует обоснованность разделения студентов по образовательным результатам на шесть групп.

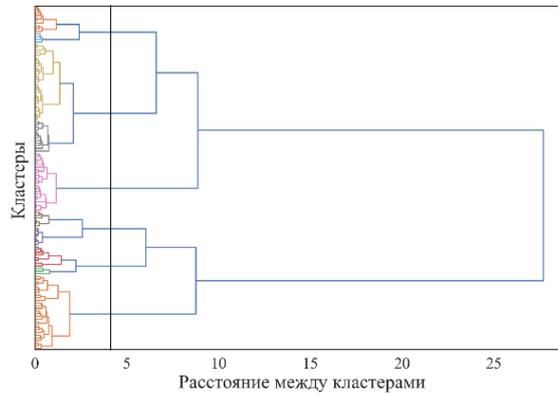


Рис. 2. Дендрограмма

В таблице 7 приведены средние характеристики кластеров. Количество объектов по кластерам распределено неравномерно: изменяется от 11 до 41. Студенты, которые сдали экзамен по дисциплине и студенты, которые его не сдали, образуют по три кластера. Среди тех, кто сдал экзамен, средний уровень посещаемости заметно выше. При этом эффективность на занятиях у них также более высокая, о чем свидетельствуют отношения $AcademF/AttendF$, которые имеют значения около 90%. Посещаемость в более поздних контрольных срезах уменьшается, а доля набранных баллов, напротив, увеличивается. В кластерах, в которых студенты сдали экзамен, средний результат на вступительных испытаниях выше, чем в кластерах со студентами, которые не сдали экзамен по дисциплине.

Таблица 7. Средние показатели кластеров

Номер кластера	Число объектов в кластере	Exam	Group	Sex	Entrance_exam	Первый РК		Второй РК		Итог за первый семестр		
						Academ1	Attend1	Academ2	Attend2	AcademF	AttendF	AcademF/AttendF, %
1	28	0	0,000	1,000	183,2	0.423	0.823	0.393	0.662	0.386	0.534	72,2
2	11	0	1,000	0,727	197,8	0.500	0.828	0.408	0.705	0.479	0.655	73,1
3	13	0	0,461	1,000	186,2	0.061	0.152	0.048	0.099	0.026	0.071	36,0
4	22	1	0,000	1,000	189,3	0.601	0.961	0.666	0.913	0.786	0.888	88,6
5	41	1	1,000	1,000	200,4	0.659	0.924	0.675	0.916	0.801	0.886	90,3
6	14	1	0,714	0,000	204,2	0.639	0.957	0.610	0.893	0.796	0.852	93,5

Первый кластер (рис. 3), среди образовавших академическую задолженность студентов, – самый многочисленный и представлен только направлением подготовки «Информатика и вычислительная техника».

Второй кластер – это студенты-задолжники, которые обучались по направлению «Программная инженерия». Доля итоговой успеваемости у второго кластера выше, чем у первого, при одинаковой эффективности работы на занятиях. Третий кластер – студенты с

крайне низкой посещаемостью и академической успеваемостью, которые не проявляли интереса к изучению дисциплины.

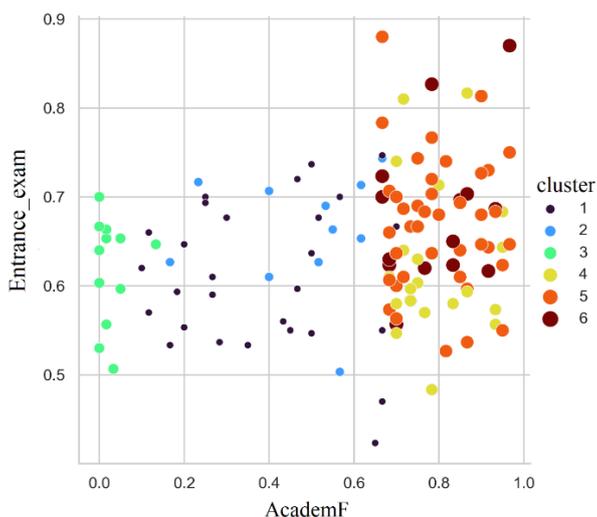


Рис. 3. Кластеры в координатах *AcademF*, *Entrance_exam*

Доля студенток, сдавших экзамен, составляет $14/17 \approx 76,5\%$, что заметно выше, чем число сдавших экзамен студентов ($63/112 \approx 56,2\%$). Одним из факторов данного различия, возможно, является тот факт, что студентки первого курса имели на момент поступления в ВУЗ более высокий средний балл вступительных испытаний. Кроме того, как правило, девушки более ответственно подходят к выполнению заданий в рамках изучаемого курса. Однако для проверки статистической гипотезы о равенстве долей данных недостаточно.

На рис. 3 показана диаграмма, визуализирующая кластеры в координатах *Entrance_exam* и *AcademF*. Из данной диаграммы видно, что разброс значений *Entrance_exam* для студентов, сдавших экзамен, выше, чем у студентов, которые экзамен не сдали. При итоговой успеваемости выше 0,667 (т.е. от 40 баллов и выше) почти отсутствуют представители кластеров с академической задолженностью. Это обусловлено тем, что студенты, которые к дате проведения экзамена не наберут хотя бы 40 баллов из 60 возможных и не выполнят все установленные учебные задания, к экзамену не допускаются и у них образуется академическая задолженность. Студенты, которые набрали 40 баллов и выше, были допущены к экзамену, хорошо осваивают учебную программу и, как правило, успешно сдают экзамен.

Заключение. Основным результатом работы является тот факт, что была показана возможность прогнозирования сдачи экзамена по отдельной дисциплине студентов первого курса вуза на основе данных контрольных срезов, которые проводят деканаты для выявления групп студентов с повышенным риском возникновения академической задолженности. Дополнительным фактором, который использовался для построения прогнозных моделей, выступал результат вступительных испытаний в Институт (ЕГЭ или внутренний экзамен). Точность прогнозирования, которую показали рассматриваемые модели (наивный байесовский классификатор и логистическая регрессия) оказалась вполне приемлемой как на этапе проведения первого рубежного контроля, так и на этапе второго.

Результаты данной работы имеют практическое значение для администрации ВУЗов и для преподавателей. Прогнозные модели можно использовать при выявлении группы студентов, имеющих высокий риск отчисления вследствие академической неуспеваемости. Модели могут быть встроены в образовательные информационные системы и быть помощником преподавателям для принятия решений в процессе реализации дисциплины.

Кластерный анализ, проведенный с использованием иерархического подхода, показал, что в кластерах, в которых студенты не сдали экзамен по дисциплине, существенно ниже уровень посещаемости студентов. Кроме того, студенты, допустившие дефолт (т.е. образование

академической задолженности по дисциплине) менее эффективно работали на учебных занятиях в течение семестра.

Список источников

1. Горбунова Е.В. Построение модели выбытия студентов по данным университетов с разной периодичностью рубежного контроля / Е.В. Горбунова, В.В. Ульянов, К.К. Фурманов // Прикладная эконометрика, 2017. – Т. 45. – С. 116-135.
2. Гафаров Ф.М. Прогностическое моделирование в высшем образовании: определение факторов академической успеваемости / Ф.М. Гафаров, Я.Б. Руднева, У.Ю. Шарифов // Высшее образование в России, 2023. – Т. 32. – №1. – С. 51-70. – DOI: 10.31992/0869-3617-2023-32-1-51-70.
3. Czibula G., Mihai A., Crivei L. S PRAR: A novel relational association rule mining classification model applied for academic performance prediction. Procedia computer science, 2019, no. 159, pp. 20-29, DOI: 10.1016/j.procs.2019.09.156.
4. Есин Р. В. Прогнозирование успешности обучения по дисциплине на основе универсальных показателей цифрового следа LMS Moodle. / Р. В. Есин, Т. А. Кустицкая, М. В. Носков // Информатика и образование, 2023. – № 38(3). – С. 31-41. – DOI: 10.32517/0234-0453-2023-38-3-31-41.
5. Русаков С.В. Нейросетевая модель прогнозирования группы риска по успеваемости студентов первого курса / С.В. Русаков, О.Л. Русакова, К.А. Посохина // Современные информационные технологии и ИТ-образование, 2018. – Т. 14. – № 4. – С. 815-822. – DOI: 10.25559/SITITO.14.201804.815-822.
6. Шухман А.Е. Анализ и прогнозирование успеваемости обучающихся при использовании цифровой образовательной среды / А.Е. Шухман, Д.И. Парфенов, Л.В. Легашев, Л.С. Гришина // Высшее образование в России, 2021. – Т. 30. – № 8-9. – С. 125-133. – DOI: 10.31992/0869-3617-2021-30-8-9-125-13.
7. Егорова Е.С. Data Mining в образовании: прогнозирование успеваемости учащихся / Е.С. Егорова, Н.А. Попова // Моделирование, оптимизация и информационные технологии, 2023. – №11(2). – DOI: 10.26102/2310-6018/2023.41.2.003. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1325> (дата обращения: 20.08.2023).
8. Шармин В. Г. Определение степени влияния различных факторов на академическую успеваемость студентов на основе их самооценки, в том числе с учетом пола студента / В. Г. Шармин, Т. Н. Шармина, Д. В. Шармин // Science for Education Today, 2022. – Т. 12. – № 3. – С. 92–114. – DOI: 10.15293/2658-6762.2203.05.
9. Шмарихина Е.С. Исследование факторов успеваемости обучающихся / Е.С. Шмарихина // Вестник НГУЭУ, 2018. – № 3. – С. 130-143.
10. Куприянов Р.Б. Повышение качества модели прогнозирования образовательных результатов студентов университетов / Р.Б. Куприянов, Д. Ю. Звонарев // Информатика и образование, 2021. – Т. 36(9). – С. 40–46. – DOI: 10.32517/0234-0453-2021-36-9-40-46.
11. Куприянов Р. Б. Разработка модели прогнозирования образовательных результатов обучающихся для университетов / Р.Б. Куприянов, Д.Ю. Звонарев // Искусственный интеллект и принятие решений, 2021. – №2. – С. 11-20. – DOI: 10.14357/20718594210202.
12. Носков М.В. Прогностическая модель оценки успешности предметного обучения в условиях цифровизации образования / М.В. Носков, Ю.В. Вайнштейн, М.В. Сомова, И.М. Федотова // Вестник Российского университета дружбы народов. Серия: Информатизация образования, 2023. – Т. 20. – № 1. – С. 7–19. – DOI:10.22363/2312-8631-2023-20-1-7-19.
13. Накарякова Н.Н. Прогнозирование группы риска (по успеваемости) среди студентов первого курса с помощью дерева решений / Н.Н. Накарякова, С.В. Русаков, О.Л. Русакова // Прикладная математика и вопросы управления. – 2020. – № 4. – С. 121-136. – DOI: 10.15593/2499-9873/2020.4.08.
14. Ерохина Е.А. Влияние результатов ЕГЭ на успеваемость студентов ВУЗ / Е.А. Ерохина, Д.В. Хруслова // Информационные технологии в науке, образовании и управлении. – Москва, 2016. – С. 265-272. – URL: https://elibrary.ru/download/elibrary_26377412_40315072.pdf.
15. Surbhi A., Santosh K. Vishwakarma, Akhilesh K. Sharma Using Data Mining classifier for predicting student's performance in UG level. International journal of Computer applications, 2017, v. 172, no.8, pp. 39-44, available at: https://www.researchgate.net/publication/319172745_Using_Data_Mining_Classifier_for_Predicting_Student's_Performance_in_UG_Level (accessed: 08/20/2023).

Алпатов Алексей Викторович. К. ф.-м. н., доцент, доцент кафедры Информатика и технология программирования, Волжский политехнический институт (филиал). Научные интересы: Моделирование социально-экономических процессов на основе эконометрического подхода и методов машинного обучения. AuthorID: 148251, ORCID: 0000-0002-3344-5984, alpatov80@mail.ru, 404121, Волгоградская область, Волжский, ул. Энгельса, 42а.

UDC 004.89+004.048

DOI: 10.25729/ESI.2023.32.4.006

Application of machine learning to analyze academic performance of university students

Aleksey V. Alpatov

Volzhsky polytechnic institute (branch) of Volgograd state technical university, Russian Federation, Volzhskiy, alpatov80@mail.ru

Abstract. The paper presents the results of analyzing and predicting the educational results of first-year university students in the implementation of a separate discipline using machine learning. The relevance of the research topic is due to the need for universities in modern conditions to successfully compete in the educational services market, which is characterized by a low number of applicants and an increase in requirements for the quality of vocational education both on the part of applicants and on the part of the state. An important component for effective decision-making in the process of quality management of the educational process is educational analytics, on the basis of which it is possible to predict the academic performance of students, to identify factors that have a significant impact on achieving high educational results. The study showed the possibility of predicting the exam in a particular discipline of first-year university students based on the data of control sections conducted by deans during the semester to identify groups of students with an increased risk of academic debt. The prediction accuracy shown by the constructed models (neural network, decision tree and logistic regression) turned out to be quite acceptable both at the stage of the first boundary control and at the stage of the second. The results of this work are of practical importance for the administration of universities and for teachers. Predictive models can be used to predict the expulsion of students due to academic failure. Models can be embedded in educational information systems and be an assistant to teachers for decision-making in the process of implementing the discipline.

Keywords: students' performance prediction, learning analytics, educational data mining, unified state exam, neural networks, decision tree, logistic regression, clustering

References

1. Gorbunova E.V., Ulyanov V.V., Furmanov K.K. Postroenie modeli vybytija studentov po dannym universitetov s raznoj periodichnost'ju rubezhnogo kontrolja [Using data from universities with different structure of academic year to model student attrition]. *Prikladnaja jekonometrika [Applied Econometrics]*, 2017, v. 45, pp. 116-135.
2. Gafarov F.M., Rudneva Ya.B., Sharifov U.Yu. Prognosticheskoe modelirovanie v vysshem obrazovanii: opredelenie faktorov akademicheskoy uspevaemosti [Predictive modeling in higher education: determining factors of academic performance]. *Vysshee obrazovanie v Rossii [Higher education in Russia]*, 2023, v. 32, no. 1, pp. 51-70, DOI: 10.31992/0869-3617-2023-32-1-51-70.
3. Czibula G., Mihai A., Crivei L. S PRAR: A novel relational association rule mining classification model applied for academic performance prediction. *Procedia computer science*, 2019, no 159, pp. 20-29, DOI: 10.1016/j.procs.2019.09.156.
4. Esin R.V., Kustitskaya T.A., Noskov M.V. Prognozirovanie uspehnosti obuchenija po discipline na osnove universal'nyh pokazatelej cifrovogo sleda LMS Moodle [Predicting academic performance in a course by universal features of LMS Moodle digital footprint]. *Informatika i obrazovanie [Informatics and Education]*, 2023, v. 38, no. 3, pp. 31-41, DOI: 10.32517/0234-0453-2023-38-3-31-41.
5. Rusakov S.V., Rusakova O.L., Posokhina K.A. Nejrosetevaja model' prognozirovanija gruppy riska po uspevaemosti studentov pervogo kursa [Neural network model of predicting the risk group for the accession of students of the first course]. *Sovremennye informacionnye tehnologii i IT-obrazovanie [Modern Information Technologies and IT-Education]*, 2018, v. 14, no 4, pp. 815-822, DOI: 10.25559/SITITO.14.201804.815-822.
6. Shukhman A.E., Parfenov D.I., Legashev L.V., Grishina L.S. Analiz i prognozirovanie uspevaemosti obuchajushihhsja pri ispol'zovanii cifrovoj obrazovatel'noj sredy [Analysis and forecasting students' academic performance using a digital educational environment]. *Vysshee obrzovanie v Rossii [Higher education in Russia]*, 2021, v. 30, no. 8-9, pp. 125-133, DOI: 10.31992/0869-3617-2021-30-8-9-125-133.
7. Egorova E.S., Popova N.A. Data Mining v obrazovanii: prognozirovanie uspevaemosti uhashhihsja [Data Mining in education: predicting student performance]. *Modelirovanie, optimizacija i informacionnye tehnologii [Modeling, Optimization and Information Technology]*, 2023, v. 11, no. 2, DOI: 10.26102/2310-6018/2023.41.2.003, available at: <https://moitvivi.ru/ru/journal/pdf?id=1325> (accessed: 08/20/2023).
8. Sharmin V.G., Sharmina T.N., Sharmin D.V. Opredelenie stepeni vlijanija razlichnyh faktorov na akademicheskuyu uspevaemost' studentov na osnove ih samoocenki, v tom chisle s uchedom pola studenta [Identifying the degree of influence of various factors on students' academic performance based on their self-assessment, taking

- into account students' gender]. Science for education today, 2022, v. 12, no. 3, pp. 92-114, DOI: 10.15293/2658-6762.2203.05.
9. Shmarikhina E.S. Issledovanie faktorov uspevaemosti obuchajushhihsja [Investigation the factors of students performance]. Vestnik NGUJeU [Vestnik NSUEM], 2018, v. 3, pp. 130-143.
 10. Kupriyanov R.B., Zvonarev D.Yu. Povyshenie kachestva modeli prognozirovaniya obrazovatel'nyh rezul'tatov studentov universitetov [Improving the quality of the university students' academic performance prediction model]. Informatika i obrazovanie [Informatics and education], 2021, v. 36, no. 9, pp. 40-46, DOI: 10.32517/0234-0453-2021-36-9-40-46.
 11. Kupriyanov R.B., Zvonarev D.Yu. Razrabotka modeli prognozirovaniya obrazovatel'nyh rezul'tatov obuchajushhihsja dlja universitetov [Developing of the student's educational success prediction model for universities]. Iskusstvennyy intellekt i prinjatie reshenij [Artificial intelligence and decision making], 2021, v.2, pp. 11-20, DOI: 10.14357/20718594210202.
 12. Noskov M.V, Vaynshteyn Yu.V, Somova M.V, Fedotova I.M. Prognosticheskaja model' ocenki uspehnosti predmetnogo obuchenija v uslovijah cifrovizacii obrazovanija [Prognostic model for assessing the success of subject learning in conditions of digitalization of education]. Vestnik Rossijskogo universiteta družby narodov. Serija: Informatizacija obrazovanija [RUDN Journal of informatization in education], 2023, vol. 20(1), pp.7-19, DOI: 10.22363/2312-8631-2023-20-1-7-19.
 13. Nakaryakova N.N., Rusakov S.V., Rusakova O.L. Prognozirovanie gruppy riska (po uspevaemosti) sredi studentov pervogo kursa s pomoshh'ju dereva reshenij [Prediction of the risk group (by academic performance) among first course students by using decision tree method]. Prikladnaja matematika i voprosy upravlenija [Applied mathematics and control sciences], 2020, no. 4, pp. 121-136, DOI: 10.15593/2499-9873/2020.4.08.
 14. Erokhina E.A., Khruslova D.K. Vliyaniye rezul'tatov YEGE na uspevyemost' studentov VU [Influence of results of use on progress of students higher education institution]. Informacionnye tehnologii v nauke, obrazovanii i upravlenii [Information technologies in science, education and management], 2016, p., 265-272, available at: https://elibrary.ru/download/elibrary_26377412_40315072.pdf.
 15. Surbhi A., Santosh K. Vishwakarma, Akhilesh K. Sharma. Using Data Mining Classifier for Predicting Student's Performance in UG Level. International journal of Computer applications, 2017, v. 172, no.8, pp. 39-44, available at: https://www.researchgate.net/publication/319172745_Using_Data_Mining_Classifier_for_Predicting_Student's_Performance_in_UG_Level. (accessed:08/20/2023).

Alexey Viktorovich Alpatov. Candidate of physics and mathematics sciences, associate professor of the department of computer science and programming technology, Volzhsky Polytechnic Institute (branch) of Volgograd state technical university. Research interests: Modeling of socio-economic systems based on the econometric approach and machine learning. AuthorID (RSCI): 148251, Researcher ID(Scopus): 57170357100, ORCID: 0000-0002-3344-5984, alpatov80@mail.ru, Russia, Volgograd region, Volzhskiy, Engels str., 42a.

Статья поступила в редакцию 06.12.2023; одобрена после рецензирования 13.12.2023; принята к публикации 14.12.2023.

The article was submitted 12/06/2023; approved after reviewing 13/12/2023; accepted for publication 12/14/2023.

Математическое моделирование в энергетике

УДК 519-7

DOI:10.25729/ESI.2023.32.4.007

Математическое моделирование чувствительности собственных колебаний конструкций с учетом вращения и неравномерного нагрева

Репецкий Олег Владимирович, Хоанг Динь Кыонг

Иркутский государственный аграрный университет им. А.А. Ежевского,

Россия, Иркутск, *hoangcuonghd95@gmail.com*

Аннотация. В данной работе разработаны математические модели и проведено тестирование чувствительности колебаний пластины с учетом вращения, неравномерного нагрева с распространением численного анализа на радиальные рабочие колеса турбомашин. Сначала, исследовалось влияние линейного и квадратичного закона изменения температуры на статические и динамические характеристики пластины с помощью программ BLADIS + и ANSYS. Эти результаты расчета собственных частот согласованы в двух программах и с аналитическим решением. В процессе расчета отмечается, что собственные частоты пластины уменьшаются с квадратичным законом изменения температуры, но увеличиваются с учетом вращения. Для понимания влияния вращения и неравномерного нагрева на чувствительность с помощью программ SOLIDWORKS, ANSYS WORKBENCH и MATLAB моделируются режимы вибрации и их чувствительность к изменению исследуемых частот. Принято на основании результатов тестирования пластины, которые показали хорошую сходимость расчета собственных частот, распространить созданные математические модели и комплексы программ на радиальные рабочие колеса реальных турбомашин. Данный покрывной диск колеса характеризуется наибольшей степенью деформации с учетом вращения, а зона снижения частот колебаний уменьшается на верхней кромке лопаток, но увеличивается в середине входной кромки лопаток с учетом вращения и неравномерного нагрева. Таким образом, полученные расчеты чувствительности колебаний пластины и реального рабочего радиального колеса от изменения массы в узлах конечно-элементной модели с учетом вращения и неравномерного нагрева позволяют сократить объем дорогостоящих экспериментальных исследований и уменьшить сроки конструирования новых машин.

Ключевые слова: конечно-элементная модель, надежность, пластина, радиальное рабочее колесо, турбомашин, частота колебаний, чувствительность.

Цитирование: Репецкий О.В. Математическое моделирование чувствительности собственных колебаний конструкций с учетом вращения и неравномерного нагрева / О.В. Репецкий, Д.К. Хоанг // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 79-90. – DOI:10.25729/ESI.2023.32.4.007.

Введение. В реальных условиях примерно 60% отказов двигателей в эксплуатации связано с разрушением или поломкой деталей ротора из-за их недостаточной прочности. Поврежденными деталями часто являются высоконагруженные элементы радиальных роторов, а именно, рабочие лопатки, в которых накапливается повреждаемость от влияния центробежных сил, температурного нагрева и аэродинамического влияния [1, 2, 3]. Из этого количества поломок примерно 70% деталей разрушается вследствие их вибраций. Задача обеспечения высокой прочности и долговечности радиальных рабочих лопаток с разнообразием их геометрических форм требует дальнейшего анализа.

Исследования разных авторов показали, что анализ чувствительности собственных колебаний роторных конструкций – это путь к улучшению прочностных характеристик энергетических турбомашин [1, 4-6], так как результаты этих исследований позволяют целенаправленно влиять на их вибрационный спектр. Исследование чувствительности является очень полезным инструментом в ряде численных процедур, таких, как идентификация параметров, обновление модели, оптимальное проектирование и др. [4, 7, 8]. Анализ чувствительности является известным численным подходом, предназначенным для исследования упругих конструкций, вычисляемых на основе аналитических частотных характеристик по отношению к набору конструктивных параметров [5, 9]. Комплексное

исследование чувствительности собственных колебаний конструкций с учетом эксплуатационных факторов дает достаточный прогноз и точность при проектировании новых турбомашин для авиационного, химического, транспортного и энергетического машиностроения [6].

Анализ чувствительности конструкций часто применяется и исследуется на стадии проектирования конструкций турбин. Используя расчеты чувствительности, разработчики заранее имитируют различные варианты проектирования элементов, удовлетворяющие всем требованиям. Эти расчеты также позволяют многократно изменять конструкцию и проводить проверочный расчет на прочность до окончательного получения завершения проекта [6, 10]. Анализ чувствительности также позволяет создать эффективные модели преднамеренной расстройки, снизить уровни возникающих напряжений и увеличить долговечность машин [11-13]. Использование математических моделей для анализа чувствительности позволяет целенаправленно устанавливать расстройку параметров, сократить объем дорогостоящих экспериментальных исследований, уменьшить сроки конструирования новых машин по критериям эффективности, надежности и ресурсосбережения их высоконагруженных агрегатов [2].

Наибольшее развитие анализа чувствительности получило в работах Репецкого О.В., Костюк А.Г., Хог Э., Чой К. и их соавторов [1, 5-7, 10, 14]. Анализируя общие результаты этих исследований, полагаем, что наиболее универсальным и широко применяемым на практике методом расчета механических систем является метод конечных элементов (МКЭ). Также следует отметить, что фактически отсутствует анализ чувствительности собственных колебаний конструкций роторов энергетических турбомашин с учетом вращения и неравномерного нагрева. В этой связи, разработка математических моделей для исследования чувствительности колебаний роторов энергетических турбомашин с учетом вращения, и неравномерного нагрева является актуальной научной задачей, требующей своего решения.

Таким образом, в данной публикации разработаны математические модели и проведено тестирование чувствительности колебаний пластины с учетом вращения, и неравномерного нагрева с распространением численного анализа на радиальные рабочие колеса турбомашин. С помощью программы SOLIDWORKS, ANSYS WORKBENCH и MATLAB моделируются режимы вибрации, и анализируется их чувствительность к изменению исследуемых частот. Результаты показывают, что собственные частоты чувствительны к изменению степени деформации.

Собственные частоты колебаний вычисляются из уравнения [3, 15]:

$$M\ddot{\delta} + K\delta = 0, \quad (1)$$

где δ – вектор перемещений, K и M – матрицы жесткости и массы.

Исправленный проект при анализе чувствительности собственных частот колебаний имеет параметры:

$$X_1 = X_0 + \delta X. \quad (2)$$

где X_0 – исходный вариант, X_1 – измененный вариант исследования.

Выражение для вычисления чувствительности частот собственных колебаний имеет вид:

$$\frac{df}{dX_i} = \frac{df}{d\lambda} \cdot \frac{d\lambda}{dX_i} = \frac{1}{4\pi\sqrt{\lambda}} \cdot \frac{d\lambda}{dX_i}, \quad (3)$$

где $\lambda = (2\pi f)^2$ – собственное значение.

Для определения зависимости расположения зон чувствительности частот колебаний от распределения масс предложена математическая зависимость вида:

$$\Delta f_k^{\max} = \frac{f_k^{\max} - f_0}{f_0}, \quad \Delta f_k^{\min} = \frac{f_k^{\min} - f_0}{f_0}, \quad (4)$$

где Δf_k^{\min} и Δf_k^{\max} – минимальное и максимальное значение чувствительности собственной частоты; f_0, f_k – значение собственных частот колебаний лопатки без учета и с учетом влияния вращения и неравномерного нагрева в каждом k -ом узле конечного элемента.

1. Тестирование пластины на влияние температуры и вращения. Как показывают исследования многих авторов, метод конечных элементов (МКЭ) предлагается как приближенный метод решения задач на прочность и долговечность роторов турбомашин. МКЭ имеет достаточно общий алгоритм, который позволяет быстро выполнить расчеты различных вариантов сложных конструкций, и прост в использовании [16-18]. МКЭ удачно конкурирует с методом граничных элементов за счет преимуществ при анализе конструкций сложной геометрии и задач, где почти каждый элемент исследуемой среды обладает различными свойствами. В данной работе МКЭ применен для решения задачи тестирования колебаний пластины на вращение и температуру.

Собственные частоты колебаний с учетом вращения и неравномерного нагрева в статическом НДС вычисляются из уравнения [3, 6]:

$$(K_E + K_G + K_R)\delta = F_\Omega + F_T + F_G, \quad (5)$$

а для свободной вибрации:

$$M\ddot{\delta} + C\dot{\delta} + (K_E + K_G + K_R)\delta = 0, \quad (6)$$

где δ – вектор перемещений, K_E и M – матрицы жесткости и массы, соответственно, K_G – матрица геометрической жесткости, зависящая от скорости вращения и температуры, K_R – дополнительная матрица жесткости (псевдомасса), возникающая в результате вращения, F_Ω, F_T, F_G – векторы, соответствующие силам вращения, температуры и давления газа соответственно, C – матрица демпфирования.

Для тестирования алгоритма учета неравномерного нагрева рассмотрена задача нагружения консольной пластины температурой. Длина пластины – 0.069 м, ширина – 0.0175 м, толщина – 0.004 м, модуль упругости материала – $2.0 \cdot 10^5$ МПа, плотность – $7.85 \cdot 10^3$ кг/м³, коэффициент Пуассона – 0.3 [19]. Исследовалось влияние квадратичного закона изменения температуры на статические и динамические характеристики пластины (рис.1). По высоте пластины выдерживался квадратичный закон распределения температуры.

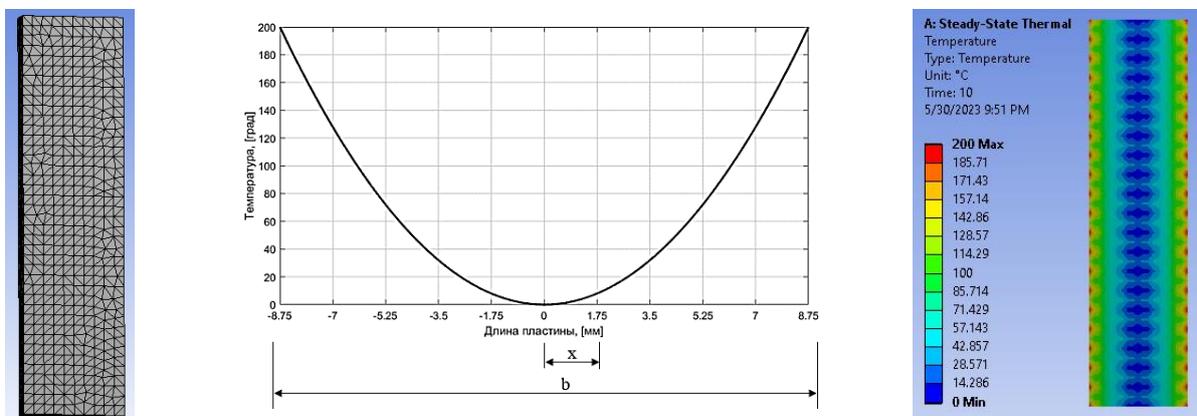


Рис. 1. Распределение температуры по хорде пластины

Квадратичный закон изменения температуры из работы [19] имеет вид:

$$T = T_{\max} x^2 / (b/2)^2, \quad (7)$$

где T – температура нагрева в конкретной точке, T_{\max} – максимальная температура нагрева при тестировании, x – расстояние от середины до края пластины, b – ширина (хорда) пластины.

В таблице 1 исследовалось влияние линейного и квадратичного закона изменения температуры на собственные частоты колебаний пластины. Видно, что собственные частоты с учетом температуры пластины неплохо согласуются в программах BLADIS + [20] и ANSYS. Собственные частоты с учетом квадратичного закона изменения температуры уменьшаются в сравнении с $T = 20^{\circ}\text{C}$.

Таблица 1. Собственные частоты с учетом нагрева пластины

Форма колебаний	Температура нагрева	Собственные частоты, Гц		
		BLADIS +	ANSYS	Отклонение, %
1И	$T = 20^{\circ}\text{C}$	635.8	638.73	+0.46
	$T = T_{\max} x^2 / (b/2)^2$	578.6	579.31	+0.12
2И	$T = 20^{\circ}\text{C}$	4505.5	4480.8	-0.55
	$T = T_{\max} x^2 / (b/2)^2$	4459.2	4478.2	+0.43
3И	$T = 20^{\circ}\text{C}$	13072.2	13086	+0.11
	$T = T_{\max} x^2 / (b/2)^2$	13025.3	13080	+0.42
1К	$T = 20^{\circ}\text{C}$	5663.9	5646.7	-0.30
	$T = T_{\max} x^2 / (b/2)^2$	5468.2	5415.3	-0.97
1ИК	$T = 20^{\circ}\text{C}$	17713.4	17773	+0.34
	$T = T_{\max} x^2 / (b/2)^2$	17569.5	16870	-3.98
2ИК	$T = 20^{\circ}\text{C}$	31708.8	31800	+0.29
	$T = T_{\max} x^2 / (b/2)^2$	31557.9	31482	-0.24

Согласно классическому подходу [19], линейное изменение температуры не создает напряжений в стержне. При увеличении температуры или перепада температур между верхней и нижней поверхностями пластина увеличивает свою длину и расширяется к корневому сечению. Также наблюдается интенсивный изгиб в плоскости наименьшей жесткости.

При анализе колебаний получено, что влияние нагрева наиболее существенно сказывается на изгибных колебаниях в плоскости минимальной жесткости. В меньшей мере изменяются крутильные колебания. При изгибно-крутильных колебаниях происходит незначительное снижение собственных частот. При изгибе в плоскости максимальной жесткости снижение частот не проходит. Первые формы колебаний пластины без нагрева представлены на Рис. 2.

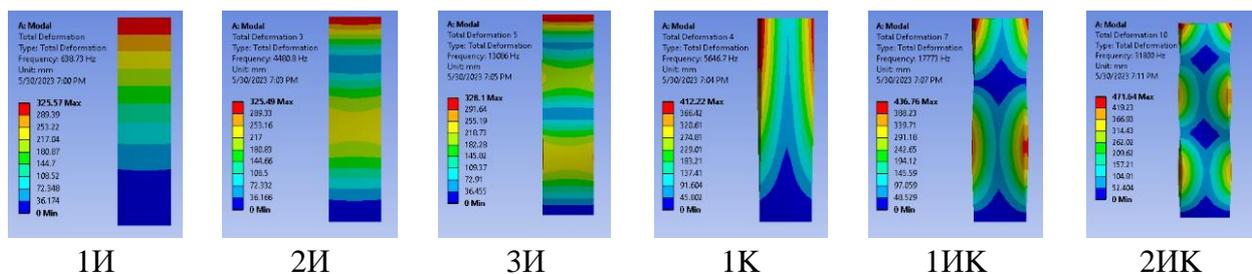


Рис. 2. Некоторые формы колебаний пластины

Для тестирования алгоритма учета вращения была рассмотрена консольная пластина постоянной толщины. Исследуемая модель пластины с жесткой заделкой в корне аппроксимировалась программой BLADIS + (КЭ STIO18), как двумерная модель [3]. В

трехмерной конечно-элементной модели применялся конечный элемент TET10 программы ANSYS WORKBENCH. Результаты расчетов по МКЭ и аналитическому решению [19, 21, 22] приведены в таблице 2. Отмечается, что собственные частоты пластины увеличиваются с учетом вращения и достаточно хорошо совпадают с аналитическим решением.

Таблица 2. Собственные частоты с учетом вращения пластины

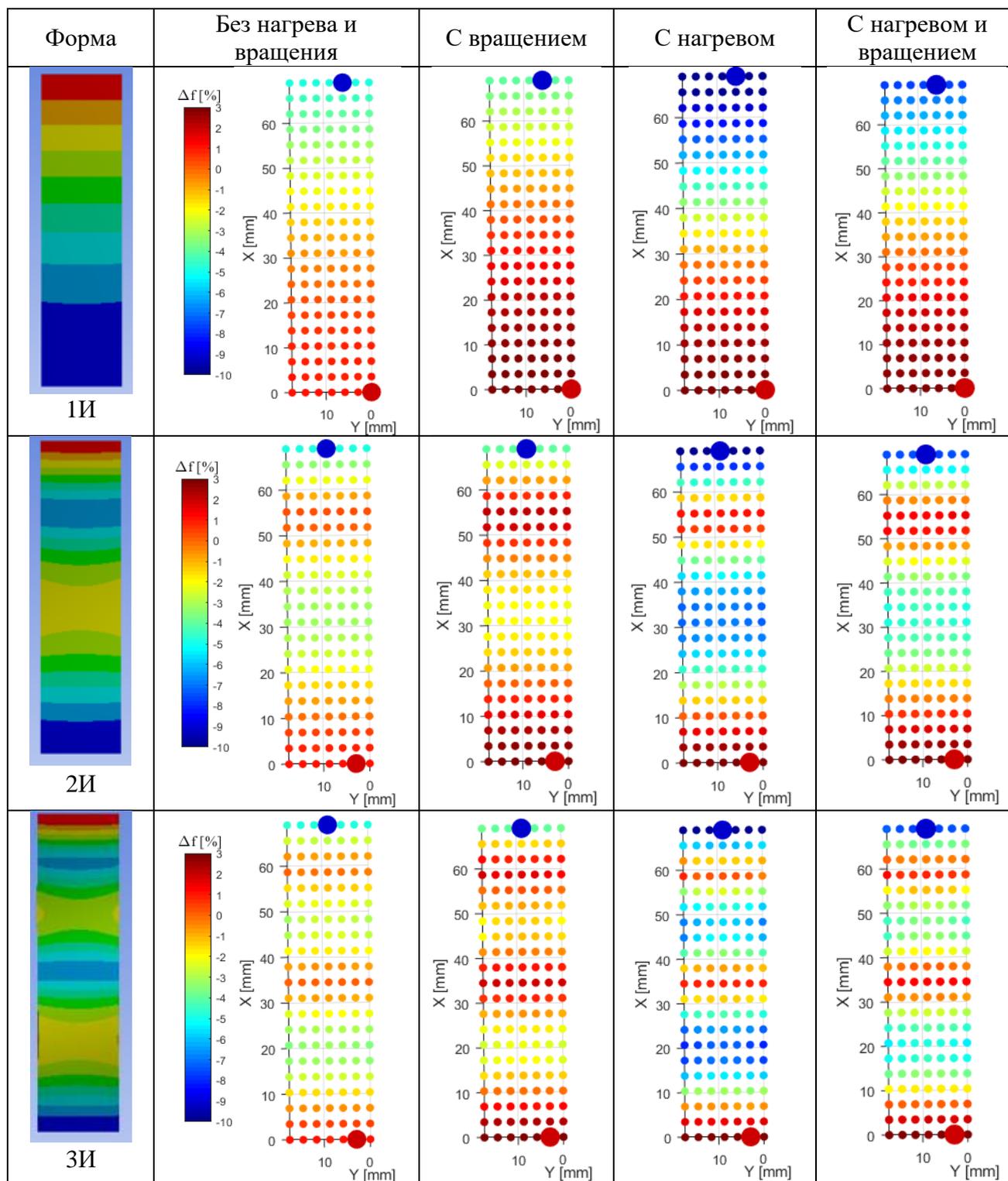
Формы колебаний	Скорость вращения пластины, рад/с						
	0	500	$\Delta f, \%$	1000	$\Delta f, \%$	1500	$\Delta f, \%$
Аналитическое решение							
1И	715	735	+2.8	792	+10.8	879	+22.9
2И	4784	4787	+0.1	4796	+0.25	4811	+0.57
1К	5612	5614	+0.05	5622	+0.18	5635	+0.42
МКЭ (BLADIS +)							
1И	715	736	+2.9	797	+11.5	889	+24.4
2И	4784	4806	+0.4	4870	+1.8	4976	+4.0
1К	5612	5617	+0.08	5615	+0.05	5631	+0.34
МКЭ (ANSYS)							
1И	714.6	726.61	+1.68	788.91	+10.40	881.29	+23.33
2И	4782.8	4792.3	+0.20	4855.4	+1.52	4903.6	+2.53
1К	5611.7	5613.6	+0.03	5616.2	+0.08	5632.8	+0.38

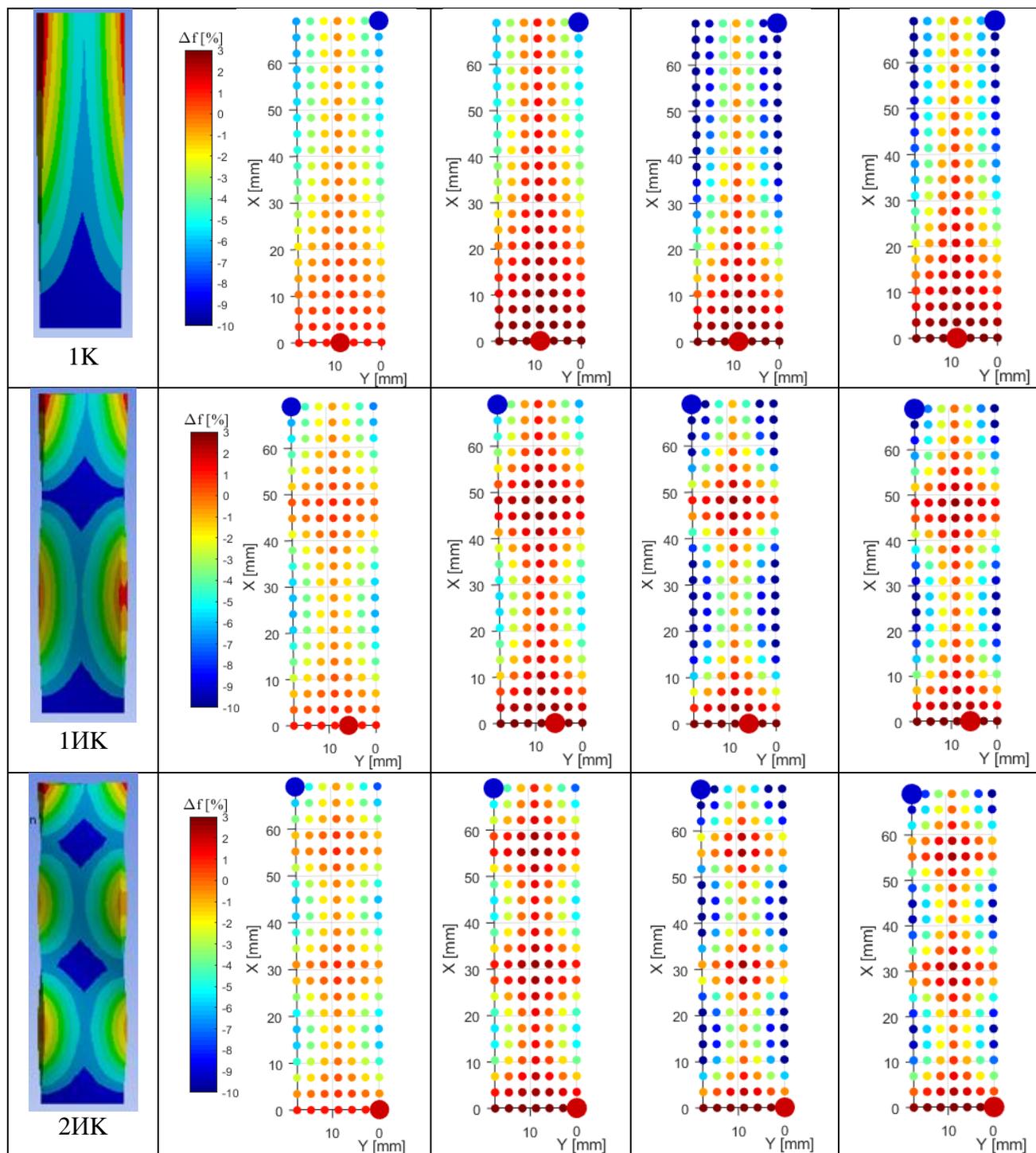
2. Анализ чувствительности пластины. Представлены шесть первых форм колебаний для анализа чувствительности пластины. Соответствующие частоты колебаний пластины приведены в таблице 1. Чувствительность просчитана и показана для пластины, где видны точки максимального и минимального изменения частот для каждой исследованной формы колебаний. Изменение цвета от красного до синего в каждом узле означает проценты отклонения частот колебаний пластины. Зона синего цвета означает максимальное снижение частоты, а красный цвет – минимальное снижение частоты колебаний. Выполнен расчет чувствительности пластины в условиях: вращение пластины $\omega = 0...500 \text{ рад/с}$, температура нагрева $T = 0...200^\circ \text{C}$ (уравнение 7).

В таблице 3 показаны результаты анализа чувствительности пластины без вращения, с вращением, с температурой и суммарным влиянием вращения и температуры. Выполненный анализ проведен как с помощью известных программных средств, так и на основе оригинальной авторской программы «Sens_Rad» [2]. Видно, что зоны распределения чувствительности пластины изменяются. Зоны повышения частот (красный цвет $\Delta f > 0$) расширяются с учетом вращения, а с учетом нагрева, расширяются зоны снижения частот (синий цвет $\Delta f < 0$). Зоны распределения чувствительности пластины комбинируются в процессе учета суммарного вращения и температуры.

Отмечается, что зоны синего цвета согласуются с зонами деформации пластины (узловыми линиями). Если зона характеризуется наибольшим снижением частот колебаний, то имеется большая степень деформации пластины. С другой стороны, зона, характеризующая наибольшим повышением частот, имеет меньшую степень деформации пластины. Таким образом, зонами снижения частот являются опасные зоны, которые передают наибольшую деформацию пластины. Согласно результатам таблицы видно, что пластина с учетом вращения имеет меньшую степень деформации (для выбранных условий), чем с учетом температурного нагрева. Однако это может измениться при других условиях нагружения.

Таблица 3. Результаты анализа чувствительности пластины





3. Анализ чувствительности покрывного диска и лопаток радиального колеса. Для понимания влияния вращения на частоту радиального колеса, в данной работе представлен расчет чувствительности покрывного диска колеса без учета и с учетом вращения. Объектом исследования является радиальное рабочее колесо с 10-ю лопатками, изготовленное фирмой “Schiele”, специализирующейся на производстве агрегатов для химической промышленности и вентиляционного оборудования. Конечно-элементная модель радиального рабочего колеса с 10-ю лопатками представлена на рис. 3 [2]. Выбрана скорость вращения колеса $\omega = 0 \dots 60$ 1/с.

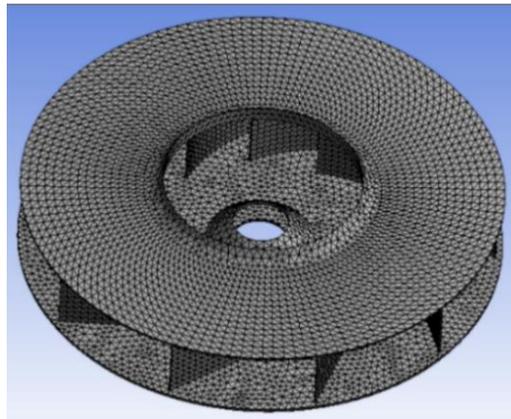
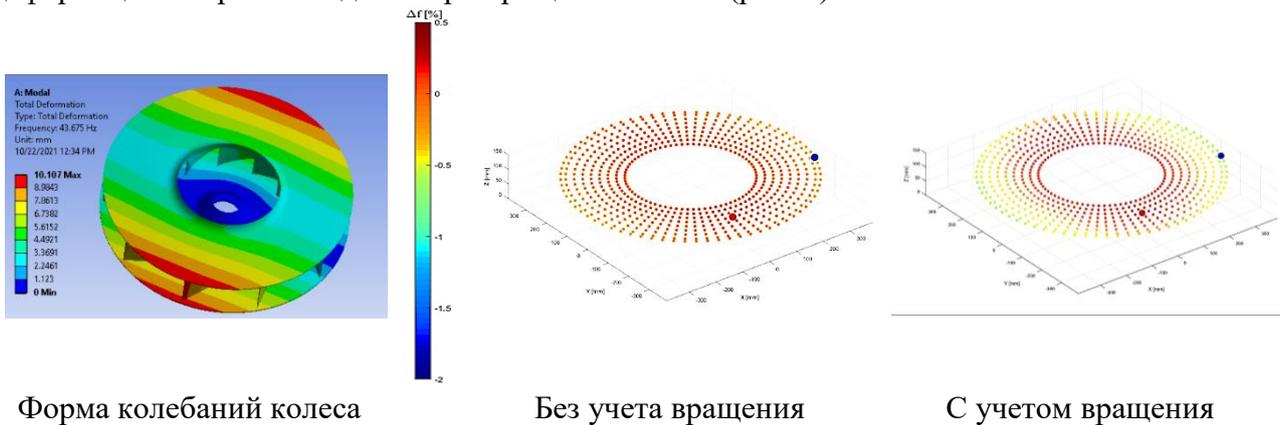


Рис. 3. Конечно-элементная модель радиального рабочего колеса с 10-ю лопатками

Материал радиального рабочего колеса – сталь, модуль Юнга – $2.1 \cdot 10^5$ МПа, плотность – 7850 кг/м^3 , коэффициент Пуассона – 0.3. Конструкция объекта жестко закреплена по ободу диска. В конечно-элементной модели применяется конечный элемент TET10 программы ANSYS WORKBENCH с общим количеством конечных элементов 58382 и 115590 узловыми точками [2, 23].

Отмечается, что зоны изменения цвета с учетом вращения покрывного диска колеса согласуются с зонами деформации покрывного диска (узловыми линиями). Если зона характеризуется наибольшим снижением частот колебаний, то имеется большая степень деформации покрывного диска при вращении колеса (рис. 4).



Форма колебаний колеса

Без учета вращения

С учетом вращения

Рис. 4. Форма колебаний и анализ чувствительности покрывного диска колеса

В процессе исследования чувствительности лопаток радиального рабочего колеса с учетом вращения $\omega = 0..60 \text{ 1/с}$ и неравномерного нагрева лопаток $T = 0..200^{\circ}\text{C}$ выявлено, что зона снижения частот колебаний (синий цвет) уменьшается на верхней кромке лопаток, но увеличивается в середине входной кромки (рис. 5). Данные результаты показывают, что середина входной кромки лопаток испытывает большую степень деформации с учетом суммарного вращения и неравномерного нагрева.

Дальнейшим этапом исследования будет анализ негативного влияния на работоспособность конструкции от зоны чувствительности данного радиального колеса с учетом вращения и неравномерного нагрева лопаток, что позволит выявить новые закономерности в чувствительности таких конструкций с учетом основных эксплуатационных факторов.

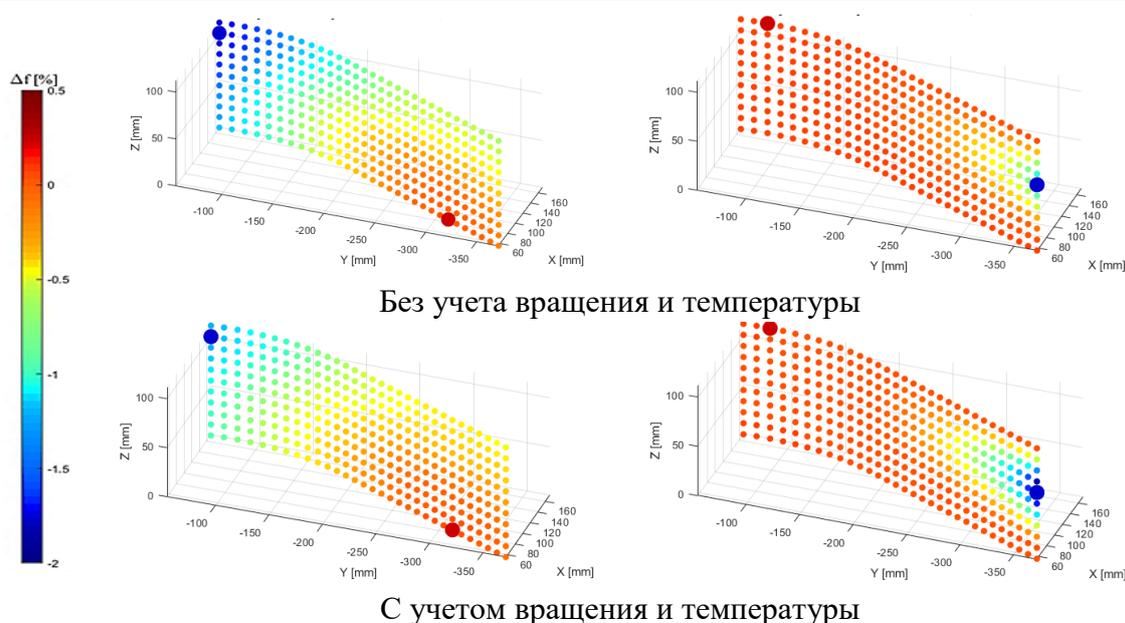


Рис. 5. Анализ чувствительности лопаток радиального рабочего колеса

Заключение. В настоящей работе разработаны математические модели и проведено тестирование чувствительности колебаний модельной пластины с учетом вращения, и неравномерного нагрева с распространением численного анализа на радиальные рабочие колеса турбомашин. Представлен расчет для тестирования собственных частот пластины в двух программах BLADIS +, ANSYS и аналитическое решение. Собственные частоты пластины прогнозируемо уменьшаются с квадратичным законом изменения температуры, но увеличиваются с учетом вращения. Проанализирована чувствительность к изменению массы основных форм колебаний пластины и покрывного диска радиального колеса с учетом вращения и неравномерного нагрева. Показаны зоны максимального и минимального изменения частот для каждой исследованной формы колебаний. Представлены зоны наибольшего снижения частот колебаний пластины, покрывного диска колеса и лопаток радиального рабочего колеса, для которых имеется большая степень деформации конструкции. Полученные результаты позволяют сократить объем дорогостоящих экспериментальных исследований радиального колеса и уменьшить сроки конструирования новых машин по критериям эффективности, надежности, технологии и ресурсосбережения высоконагруженных агрегатов машин.

Список источников

1. Костюк А.Г. Динамика и прочность турбомашин / А.Г. Костюк. – М.: Издательский дом МЭИ, 2007. – 476 с.
2. Репецкий О.В. Компьютерное моделирование и численный анализ чувствительности радиальных колес турбомашин / О.В. Репецкий, Д.К. Хоанг // Вестник НГИЭИ, 2022. – № 7 (134). – С. 22-36.
3. Репецкий О.В. Компьютерный анализ динамики и прочности турбомашин / О.В. Репецкий. – Иркутск: Изд-во ИрГТУ, 1999. – 301 с.
4. Kenyon J.A., Griffin J.H. Forced response of turbine engine bladed disks and sensitivity to harmonic mistuning. Journal of engineering for gas turbines and power, 2003, no. 1 (125), pp. 113-120.
5. Repetskii O., Zainchkovski K. The sensitivity analysis for life estimation of turbine blades. ASME 1997 Turbo asia conference, 1997, 11 p.
6. Хог Э. Анализ чувствительности при проектировании конструкций / Э. Хог, 11,1217К. Чой, В. Комков. – М.: Мир, 1988. – 428 с.
7. Костюк А.Г. Колебания в турбомашинах / А.Г. Костюк. – М.: Издательский дом МЭИ, 1961. – 201 с.
8. Tan Y., Zang C., Petrov E.P. Mistuning sensitivity and optimization for bladed disks using high-fidelity models. Mechanical systems and signal processing, 2019, vol. 124, pp. 502-523.
9. Markus W., Johannes E., Oliver V., Ralf L. Automatic numerical analyses and optimization of operating maps applied to a radial compressor. ASME turbo expo 2019: turbomachinery technical conference and exposition, GT2019-91408, V02DT46A014, 12 p.

10. Хог Э. Прикладное оптимальное проектирование: Механические системы и конструкций / Э. Хог, Я. Арора. – М.: Мир, 1983. – 480 с.
11. Beck A., Brown J.M., Kaszynski A.A., Daniel L.J. Numerical methods for calculating component modes for geometric mistuning reduced-order models. *Journal of engineering for gas turbines and power*, 2022, no. 3 (144), 9 p.
12. Besem F.M., Kielb R.E., Key N.L. Forced response sensitivity of a mistuned rotor from an embedded compressor stage. *Journal of turbomachinery*, 2016, no. 3 (138), 10 p.
13. Kaneko Y. Study on vibration characteristics of single crystal blade and directionally solidified blade. *ASME Journal of engineering for gas turbines and power*, 2011, vol. 6, pp. 931-940.
14. Репецкий О.В. Исследования влияния расстройки параметров на долговечность рабочих колес турбомашин с учетом анализа чувствительности / О.В. Репецкий, В.В. Нгуен // *Вестник НГИЭИ*, 2020. – № 10(113). – С. 5-16.
15. Когаев В.П. Расчеты деталей машин и конструкций на прочность и долговечность / В.П. Когаев, Н.А. Махутов, А.П. Гусенков. – М.: Машиностроение, 1985. – 224 с.
16. Мяченков В.И. Расчеты машиностроительных конструкций методом конечных элементов / В.И. Мяченков. – М.: Машиностроение, 1989. – 520 с.
17. Irrerier H., Repetskii O. Vibration and life estimation of rotor structures. *International federation for the promotion of mechanism and machine science (IFTOMM)*, 1998, pp. 456-464.
18. Ewins D.J. Vibration characteristics of bladed disc assemblies. *Journal of mechanical engineering science*, 1973, no. 5 (12), pp. 165-186.
19. Тимошенко С.П. Пластинки и оболочки / С.П. Тимошенко, С. Войновский-Кригер. – М.: Наука, 1965. – 635 с.
20. Repetskii O., Ryjikov I. Modeling and simulation of dynamic processes with the help of program package BLADIS+. *Innovations and advanced techniques in systems, computing sciences and software engineering*, 2008, pp. 219-220.
21. Еловенко Д.А., Экспериментальное исследование модели автоклава для гидротермального синтеза минералов / Д.А. Еловенко, П.Г. Пимштейн, О.В. Репецкий, Д.В. Татаринев // *Вестник Байкальского союза стипендиатов DAAD (Байкальский государственный университет экономики и права)*, 2010. – № 1. – С. 11-19.
22. Repetsky O.V., Nguyen T.Q., Ryzhikov I.N. Investigation of vibration and fatigue life of mistuned bladed disks. *Proceedings of the international conference actual issues of mechanical engineering 2017(AIME 2017)*, 2017, pp. 702-707
23. Repetskii O.V., Hoang D.C. Physical and mathematical modeling and computer analysis of radial impellers for chemical and power engineering, taking into account ecology. *IOP conference series: earth and environmental science*, 2022, vol. 990, 012044, 6 p.

Репецкий Олег Владимирович. Доктор технических наук, профессор, проректор по международным связям Иркутского государственного аграрного университета им. А.А. Ежовского, AuthorID: 118300, SPIN: 6232-8930, ORCID: 0000-0003-2560-2721, repetskii@igsha.ru, 664038, Россия, г. Иркутск, пос. Молодежный.

Хоанг Динь Кьонг. Аспирант кафедры Электрооборудования и физики, Иркутский государственный аграрный университет им. А.А. Ежовского, AuthorID: 1132852; SPIN: 6487-4299; ORCID: 0000-0003-0232-8723, hoangcuonghd95@gmail.com, 664038, Россия, г. Иркутск, пос. Молодежный.

UDC 519-7

DOI:10.25729/ESI.2023.32.4.007

Mathematical modeling of the sensitivity of natural vibrations of structures, taking into account rotation and heating

Oleg V. Repetskii, Dinh Cuong Hoang

Irkutsk State Agrarian University named after A.A. Ezhevsky,

Russia, Irkutsk, *hoangcuonghd95@gmail.com*

Abstract. In this work, mathematical models have been developed and the sensitivity of plate vibrations has been tested taking into account rotation, heating with the extension of numerical analysis to radial wheels of turbomachines. First, the influence of the linear and quadratic law of temperature change on the static and dynamic

characteristics of the plate was studied using the BLADIS + and ANSYS programs. These results of the calculation of natural frequencies are consistent in two programs and with an analytical solution. During the calculation, it is noted that the natural frequencies of the plate decrease with a quadratic law of temperature change, but increase with rotation. To understand the effect of rotation and heating on sensitivity using the program SOLIDWORKS, ANSYS WORKBENCH, and MATLAB simulate vibration modes and their sensitivity to changes in the studied frequencies. From the results of testing the plate, which showed good convergence of the calculation of natural frequencies, it was decided to extend the created mathematical models and software packages to the real radial wheels of turbomachines. This cover disc is characterized by the greatest degree of deformation, taking into account rotation. And the zone of reduction in vibration frequencies decreases at the upper edge of the blades, but increases in the middle of the leading edge of the blades, taking into account rotation and uneven heating. Thus, the obtained calculations of the analysis sensitivity on the plate and the real radial wheel, taking into account rotation and heating make it possible to reduce the volume of expensive experimental studies and reduce the time for designing new machines according to the criteria of efficiency, reliability, technology and resource saving of highly loaded machine units.

Keywords: finite element model, reliability, plate, radial wheel, turbomachine, frequency vibration, sensitivity

References

1. Kostyuk A.G. Dinamika i prochnost' turbomashin [Dynamics and strength of turbomachines]. Moscow, Izdatel'skiy dom MEI [Publishing House MPEI], 2007, 476 p.
2. Repetskii O.V., Hoang D.C. Komp'yuternoye modelirovaniye i chislennyy analiz chuvstvitel'nosti radial'nykh koles turbomashin [Computer simulation and numerical analysis of the sensitivity of radial wheels turbomachines]. Vestnik NGIEI, 2022, vol. 134, no. 7, pp. 22–36.
3. Repetskii O.V. Komp'yuternyy analiz dinamiki i prochnosti turbomashin [Computer analysis of the dynamics and strength of turbomachines]. Irkutsk: Izd-vo IrGTU, 1999, 301 p.
4. Kenyon J.A., Griffin J.H. Forced response of turbine engine bladed disks and sensitivity to harmonic mistuning. Journal of engineering for gas turbines and power, 2003, no. 1 (125), pp. 113- 120.
5. Repetskii O., Zainchkovski K. The sensitivity analysis for life estimation of turbine blades. ASME 1997 Turbo asia conference, 1997, 11p.
6. Khog E., Choy K., Komkov V. Analiz chuvstvitel'nosti pri proyektirovanii konstruktsiy [Sensitivity analysis in the design of structures]. Moscow, Mir, 1988, 428 p.
7. Kostyuk A. G. Kolebaniya v turbomashinakh [Vibrations in turbomachines]. Moscow, Izdatel'skiy dom MEI [Publishing House MPEI], 1961, 201 p.
8. Tan Y., Zang C., Petrov E.P. Mistuning sensitivity and optimization for bladed disks using high-fidelity models. Mechanical systems and signal processing, 2019, vol. 124, pp. 502-523.
9. Markus W., Johannes E., Oliver V., Ralf L. Automatic numerical analyses and optimization of operating maps applied to a radial compressor. ASME turbo expo 2019: turbomachinery technical conference and exposition, GT2019-91408, V02DT46A014, 12 p.
10. Khog E., Arora YA. Prikladnoye optimal'noye proyektirovaniye: Mekhanicheskiye sistemy i konstruktsiy [Applied optimal design: Mechanical systems and structures]. Moscow, Mir, 1983, 480 p.
11. Beck J.A., Brown J.M., Kaszynski A.A., Daniel L. Numerical methods for calculating component modes for geometric mistuning reduced-order models. Journal of engineering for gas turbines and power, 2022, vol. 144, no. 3, 9 p.
12. Besem F.M., Kielb R.E., Key N.L. Forced response sensitivity of a mistuned rotor from an embedded compressor stage. Journal of turbomachinery, 2016, no. 3 (138), 10 p.
13. Kaneko Y. Study on vibration characteristics of single crystal blade and directionally solidified blade. ASME Journal of engineering for gas turbines and power, 2011, vol. 6, pp. 931-940.
14. Repetskii O.V., Nguyen V.V. Issledovaniya vliyaniya rasstroyki parametrov na dolgovechnost' rabochikh koles turbomashin s uchetom analiza chuvstvitel'nosti [Study of the influence of parameter mismatch on the durability of turbomachinery impellers taking into account sensitivity analysis]. Vestnik NGIEI, 2020, no. 10(113), pp. 5-16.
15. Kogayev V.P., Makhutov N. A., Gusenkov A. P. Raschety detaley mashin i konstruktsiy na prochnost' i dolgovechnost' [Calculations of machine parts and structures for strength and durability]. Moscow, Mashinostroenie [Mechanical engineering], 1985, 224 p.
16. Myachenkov V.I. Raschety mashinostroitel'nykh konstruktsiy metodom konechnykh elementov [Calculations of machine-building structures by the finite element method]. Moscow, Mashinostroyeniye[Mechanical engineering], 1989, 520 p.
17. Irretier H., Repetskii O. Vibration and life estimation of rotor structures. International federation for the promotion of mechanism and machine science (IFTOMM), 1998, pp. 456-464.

18. Ewins D.J. Vibration characteristics of bladed disc assemblies. *Journal of mechanical engineering science*, 1973, no. 5 (12), pp. 165-186.
19. Timoshenko S.P., Voynovskiy-Kruger S. *Plastinki i obolochki [Plates and shells]*. Moscow, Nauka [Science], 1965, 635 p.
20. Repetskii O., Ryjikov I. Modeling and simulation of dynamic processes with the help of program package BLADIS+. *Innovations and advanced techniques in systems, computing sciences and software engineering*, 2008, pp. 219-220.
21. Elovenko D.A., Pimshtein P.G., Repetsky O.V., Tatarinov D.V. Eksperimental'noye issledovaniye modeli avtoklava dlya gidrotermal'nogo sinteza [Experimental study of the autoclave model for hydrothermal synthesis of minerals]. *Vestnik Baykal'skogo soyuza stipendiatov DAAD (Baykal'skiy gosudarstvennyy universitet ekonomiki i prava) [Bulletin of the Baikal Union of Scholars DAAD (Baikal State University of Economics and Law)]*, 2010, no. 1, pp. 11-19.
22. Repetsky O.V., Nguyen T.Q., Ryzhikov I.N. Investigation of vibration and fatigue life of mistuned bladed disks. *Proceedings of the international conference actual issues of mechanical engineering 2017(AIME 2017)*, 2017, pp. 702-707
23. Repetskii O.V., Hoang D.C. Physical and mathematical modeling and computer analysis of radial impellers for chemical and power engineering, taking into account ecology. *IOP conference series: earth and environmental science*, 2022, vol. 990, 012044, 6 p.

Repetskii Oleg Vladimirovich. *Doctor of technical sciences, professor, Vice-Rector for international relations, Irkutsk state Agrarian university named after A.A. Ezhevsky, AuthorID: 118300, SPIN: 6232–8930, ORCID: 0000–0003–2560–2721, repetskii@igsha.ru, 664038, Russia, Irkutsk, Molodezhny settlement.*

Dinh Cuong Hoang. *Postgraduate student of the department Electrical power and physics, Irkutsk state Agrarian university named after A.A. Ezhevsky, AuthorID: 1132852, SPIN: 6487-4299, ORCID: 0000-0003-0232-8723, hoangcuonghd95@gmail.com, 664038, Russia, Irkutsk, Molodezhny settlement.*

Статья поступила в редакцию 27.06.2023; одобрена после рецензирования 31.10.2023; принята к публикации 16.12.2023.

The article was submitted 06/27/2023; approved after reviewing 10/31/2023; accepted for publication 12/16/2023.

УДК 519.6+517.9

DOI:10.25729/ESI.2023.32.4.008

Идентификация внутреннего теплового источника и определение теплового состояния объекта по динамическим граничным измерениям

Япарова Наталья Михайловна, Капелюшин Юрий Евгеньевич

Южно-Уральский государственный университет (НИУ),

Россия, Челябинск, *iaparovant@susu.ru*

Аннотация. В статье рассмотрена задача идентификации внутреннего теплового источника и оценки его влияния на изменения температуры контролируемого объекта. Проблема идентификации теплового источника возникает при тепловом контроле передачи электроэнергии, при термообработке, при неразрушающем контроле зданий, конструкций и материалов. Математическая модель теплопереноса внутри объекта представлена уравнением теплопроводности с неизвестной функцией источника, начальными условиями, а также граничными условиями, сформированными на основе результатов зашумленных температурных измерений, полученных вблизи поверхности объекта. В статье предложен подход к идентификации внутреннего теплового источника, основанный на переходе от обратной задачи к интегральному уравнению, численный метод его решения, а также алгоритм расчета нестационарных внутренних тепловых полей, учитывающий влияние теплового источника. Устойчивость метода идентификации относительно погрешности исходных данных обеспечивается выбором параметров регуляризации. Предложенные подход и методы, в отличие от существующих, позволяют установить явную зависимость искомой функции внутреннего теплового источника от граничных измерений в ситуации, когда температурные поля вблизи поверхности объекта изменяются относительно времени. В статье приведены оценки погрешности численных решений, найденные в результате сравнительного анализа с тестовыми значениями. Результаты эксперимента свидетельствуют о том, что предложенные методы снижают негативное влияние шума на точность обработки данных и позволяют определять внутреннее тепловое состояние объекта из косвенных измерений с достаточным уровнем точности и могут служить основой для определения влияния внутреннего теплового источника на формирование внутренних нестационарных температурных полей.

Ключевые слова: теплоперенос, технические системы, метод обработки информации, обратная задача, интегральное уравнение, численный метод, регуляризация

Цитирование: Япарова Н.М. Идентификация внутреннего теплового источника и определение теплового состояния объекта по динамическим граничным измерениям / Н.М. Япарова, Ю.Е. Капелюшин // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 91-103. – DOI:10.25729/ESI.2023.32.4.008.

Введение. Требования к повышению эффективности производства и обеспечению эксплуатационной безопасности технических систем в металлургии, машиностроении и химической промышленности неразрывно связаны с высокоточным регулированием режима теплового воздействия, основанном на поступающей информации о тепловом состоянии контролируемого технического объекта. Существенным фактором, влияющим на тепловое состояние объекта, является наличие скрытого внутреннего источника тепла, появление которого связано с остаточными тепловыми процессами, происходящими внутри тела. Источником информации для идентификации внутреннего теплового источника служат результаты температурных измерений, выполненные вблизи поверхности контролируемого объекта. Задача идентификации источника, в которой внутренний источник тепла определяется по косвенным измерениям на поверхности, относится к классу обратных задач. Специфика обратных задач заключается в том, что для их решения необходимо использовать регуляризирующие алгоритмы, обеспечивающие устойчивость результатов относительно шума исходных данных и контролируемость уровня погрешности найденных решений.

Разработка и совершенствование регуляризирующих алгоритмов для решения обратных задач, связанных с реальной практикой, вызывает большой интерес у исследователей. Среди работ, посвященных этой тематике, отметим работу Вабищевича П.Н. [1] по исследованию

алгоритмов временного разделения для многомасштабных задач, Васильева В.И. [2], в которой рассмотрены задачи теплообмена с фазовым переходом, работу Лукьяненко Д.В., Борзунова А.А. и Шишленина М.А. [3], предложивших подход к решению коэффициентных обратных задач для нелинейных уравнений массопереноса. В качестве исследований, связанных с применением обратных задач на практике, отметим работы Короткого А.И. [4], Дилигенской А.Н. [5] и других исследователей [6-8]. В последние десятилетия идентификация внутреннего источника тепла в различных объектах относится к интенсивно развивающейся области исследований. Среди работ, посвященных разработке методов идентификации источника тепла, следует выделить работы E. Bazhlekova [9], S. Slimani [10], M. Rostamian [11]. В отношении применения теоретических результатов на практике упомянем работы D. Lesnic [12], Zou [13] и других исследователей [14-16].

В статье рассмотрена проблема идентификации неизвестного внутреннего теплового источника и последующего определения внутренних нестационарных температурных полей по результатам температурных измерений, выполненных вблизи поверхности объекта. Построено интегральное уравнение, связывающее функцию источника и граничные температурные измерения. Предложен регуляризирующий алгоритм расчета теплового источника, основанный на численном решении интегрального уравнения. Вычислительная схема базируется на использовании конечных интегральных сумм и многопараметрической регуляризации, обеспечивающей устойчивость вычислительной процедуры относительно погрешности исходных данных. Эффективность предложенного метода подтверждается результатами вычислительного эксперимента.

Постановка задачи. Проблема идентификации внутреннего теплового источника возникает при комплексной тепловой обработке, когда внутри технологического объекта могли сохраниться включения, отличающиеся по физическим характеристикам от общей структуры объекта и являющиеся внутренними источниками выделения или поглощения тепловой энергии. К таким включениям относятся пустоты, неоднородные включения, перегретые или непрогретые участки внутри тела. Появление внутренних источников возможно также в электропроводящих сетях, когда причиной возникновения источника являются изменения, связанные со старением проводки, ее повреждением, и другими негативными факторами. Внутренний источник тепла генерирует или поглощает тепло, а устранение его влияния требует дополнительных затрат, связанных либо с корректировкой параметров рабочего режима, либо с изменением параметров внешнего теплового воздействия. Таким образом, точность контроля теплового состояния объекта обеспечивается созданием качественных методов обработки температурных измерений, учитывающих влияние возможного теплового источника на формирование внутренних нестационарных температурных полей.

В статье рассмотрена проблема выявления и оценки влияния внутреннего теплового источника на тепловое состояние технического объекта. Основные предпосылки математической модели теплопереноса заключаются в следующем. К начальному моменту времени температурные поля одинаковы в каждой точке объекта. Поверхность объекта состоит из нескольких непересекающихся частей, одна из которых подвергается тепловому воздействию, одинаковому в каждой точке нагреваемой поверхности. К таким объектам относятся линейные токопроводящие системы, нагрев в которых происходит под действием электрического тока, кольцевые цилиндры, внутренняя поверхность которых подвергается термообработке, стенки термических печей, кристаллизаторов и других технических сооружений. Измерения температуры осуществляются на поверхности объекта. В случае линейных токопроводящих систем измерения происходят в контрольных точках. Интенсивность и продолжительность теплового воздействия такова, что значения тепловых

характеристик материала: удельная теплоемкость, теплопроводность и плотность материала остаются практически неизменными, а на поверхности тела исключаются резкие изменения температурных градиентов. В ситуации, когда внутренний источник тепла отсутствует, реализация выбранного режима теплового воздействия приводит к формированию определенного температурного поля на других участках поверхности. При наличии же внутреннего теплового источника появляются отклонения измеряемых в этой части температур от ожидаемых значений. В этой ситуации необходимо определить уровень влияния неизвестного внутреннего источника тепла на температурные поля и спрогнозировать тепловое состояние объекта с учетом найденного влияния внутреннего теплового источника.

Математическая модель теплопереноса с неизвестным внутренним источником тепла представлена следующим образом. Пусть t – текущий момент времени, L – расстояние от участка поверхности, подвергаемой тепловому воздействию до точки, в которой проводятся контрольные температурные измерения, x – расстояние от участка, подвергаемого тепловому воздействию, до текущей внутренней точки объекта, $x \in [0, L]$. Функция $u(x, t)$ соответствует температуре текущей внутренней точки x в текущий момент времени t . Математическая модель теплопереноса включает уравнение теплопроводности

$$u_t = au_{xx} + f(t), \quad x \in (0, L), \quad t \geq 0, \quad (1)$$

где a – коэффициент температуропроводности, $f(t)$ – неизвестная функция теплового источника, отражающая влияние внутреннего теплового источника на температурные поля. Тепловому состоянию объекта в начальный момент времени соответствуют начальное условие:

$$u(x, 0) = 0, \quad x \in [0, L]. \quad (2)$$

Значениям температур, измеренных на поверхности, подвергаемой непосредственному тепловому воздействию, соответствует граничное условие:

$$u(0, t) = \varphi(t), \quad t \geq 0. \quad (3)$$

Значения температур, измеренные на другом участке поверхности, служат основой для формирования граничной функции $u(L, t)$, представленной в математической модели граничным условием:

$$u(L, t) = \psi(t), \quad t \geq 0. \quad (4)$$

Основываясь на информации об интенсивности внешнего теплового воздействия и принимая во внимание теплообмен вблизи границ объекта, получаем следующее дополнительное условие в некоторой точке x_0 , расположенной в окрестности граничной точки $x = 0$:

$$u(x_0, t) = g(t), \quad t \geq 0, \quad (5)$$

где функция $g(t)$ сформирована из расчета плотности граничных тепловых потоков.

Основываясь на требованиях к интенсивности внешнего теплового воздействия, полагаем, что имеют место следующие условия $\varphi(t), \psi(t) \in C^{1+\eta}[0, T]$ для любого $T > 0$ при $\eta \in (0, 1)$ и существуют константы $S_\varphi, S_\psi, \beta_\varphi, \beta_\psi > 0$ такие, что

$$|\varphi(t)| \leq S_\varphi e^{\beta_\varphi t} \text{ и } |\psi(t)| \leq S_\psi e^{\beta_\psi t} \text{ при всех } t \in [0, T] \text{ для любого } T > 0.$$

Также существуют $S > 0$ и $\beta > 0$ такие, что имеет место $|u(x, t)| \leq Se^{\beta t}$ при всех $x \in [0, L]$ и $t \in [0, T]$ для любого $T > 0$. Кроме того, функции $\varphi(t), \psi(t), g(t)$ и $u(x, t)$ удовлетворяют условиям Дирихле в соответствующей области.

Задача определения влияния неизвестного внутреннего теплового источника заключается в отыскании функции $f(t)$ из соотношений (1) – (5).

При решении этой задачи необходимо учитывать, что результаты температурных измерений неизбежно содержат погрешности, уровень которых не превосходит некоторой величины $\delta > 0$, определяемой точностью средства измерения. В математической модели

присутствие погрешностей представлено следующим образом. Вместо действительных значений $\varphi(t), \psi(t), g(t)$ известны измеренные значения $\varphi_\delta(t), \psi_\delta(t), g_\delta(t)$ и уровень δ такие, что при любом $T > 0$ имеет место условие:

$$\max \left\{ \max_{t \in [0, T]} |\varphi(t) - \varphi_\delta(t)|, \max_{t \in [0, T]} |\psi(t) - \psi_\delta(t)|, \max_{t \in [0, T]} |g(t) - g_\delta(t)| \right\} \leq \delta.$$

Таким образом, в задаче (1) - (5) требуется, используя зашумленные исходные данные $\varphi_\delta(t), \psi_\delta(t), g_\delta(t)$, найти функцию $f_\delta(t)$, а затем рассчитать нестационарные температурные поля $u_\delta(x, t)$ в каждой внутренней точке $x \in (0, L)$ при всех $t \in (0, T)$ для любого $T > 0$.

Задача идентификации функции источника (1) - (5) относится к классу обратных задач и для построения методов идентификации внутреннего теплового источника, устойчивого относительно погрешности исходных данных, и позволяющих получать температурные значения с требуемым уровнем точности, необходимо разрабатывать алгоритмы, основанные на регуляризирующих подходах.

Построение интегрального уравнения идентификации. Построение метода идентификации функции источника $f_\delta(t)$ базируется на обобщении результатов, представленных в [16, 17]. Следуя основным идеям, изложенным в этих работах, разработка метода идентификации функции источника разделяется на два этапа. На первом этапе осуществляется переход от исходной задачи к интегральному уравнению идентификации, устанавливающему явную зависимость искомой функции от исходных данных. Следующий этап связан с разработкой регуляризованного алгоритма идентификации функции источника из полученного интегрального уравнения. Построение интегрального уравнения идентификации основывается на аналитическом решении следующей прямой задачи:

$$u_t = au_{xx} + f(t), \quad x \in (0, L), \quad t \geq 0, \quad (6)$$

$$u(x, 0) = 0, \quad x \in [0, L], \quad (7)$$

$$u(0, t) = \varphi(t), \quad u(L, t) = \psi(t), \quad t \geq 0. \quad (8)$$

Основываясь на свойствах функций $\varphi(t), \psi(t)$ и $u(x, t)$ применим для решения задачи (6) - (8) прямое преобразование Лапласа по переменной t . Пусть $\bar{u}(x, p)$ – изображение функции $u(x, t)$, а $\bar{\varphi}(p), \bar{\psi}(p), \bar{f}(p)$ – изображения функций $\varphi(t), \psi(t)$ и $f(t)$ соответственно. Тогда прямая задача (6) - (8) в изображениях Лапласа имеет вид:

$$d^2 \bar{u} - \frac{p}{a} \bar{u} = -\frac{\bar{f}}{a},$$

$$\bar{u}(0) = \bar{\varphi}, \quad \bar{u}(L) = \bar{\psi}.$$

Аналитическое решение этой задачи имеет вид:

$$\bar{u}(x, p) = \frac{\bar{f}(p)}{p} + \left(\bar{\varphi}(p) - \frac{\bar{f}(p)}{p} \right) \frac{sh\left(\sqrt{\frac{p}{a}}(L-x)\right)}{sh\sqrt{\frac{p}{a}}L} + \left(\bar{\psi}(p) - \frac{\bar{f}(p)}{p} \right) \frac{sh\left(\sqrt{\frac{p}{a}}x\right)}{sh\sqrt{\frac{p}{a}}L}, \quad (9)$$

Так как $\frac{sh\left(\sqrt{\frac{p}{a}}(L-x)\right)}{sh\sqrt{\frac{p}{a}}L}$ и $\frac{sh\left(\sqrt{\frac{p}{a}}x\right)}{sh\sqrt{\frac{p}{a}}L}$ имеют простые полюса в точках $\sqrt{\frac{p}{a}}L = \pm i\pi n$, а точка $p = 0$

является регулярной точкой функции $\frac{sh\left(\sqrt{\frac{p}{a}}x\right)}{sh\sqrt{\frac{p}{a}}L}$, то эти функции являются мероморфными.

Учитывая это свойство и принимая во внимание ограниченность этих функций на каждом замкнутом контуре, применим теорему Миттаг-Леффлера и теорему Коши [18], тогда получаем:

$$\frac{\operatorname{sh}\left(\sqrt{\frac{p}{a}}(L-x)\right)}{\operatorname{sh}\sqrt{\frac{p}{a}}L} = \frac{L-x}{L} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{\pi n(L-x)}{L}\right) \frac{p}{p + \frac{a\pi^2 n^2}{L^2}},$$

$$\frac{\operatorname{sh}\left(\sqrt{\frac{p}{a}}x\right)}{\operatorname{sh}\sqrt{\frac{p}{a}}L} = \frac{x}{L} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{\pi nx}{L}\right) \frac{p}{p + \frac{a\pi^2 n^2}{L^2}}.$$

Подставим последние соотношения в (9). Принимая во внимание сходимость рядов в этих соотношениях и опираясь на свойства сходящихся рядов, после приведения подобных слагаемых получаем, что изображение решения $\bar{u}(x, p)$ прямой задачи (6) - (8) при $x \in (0, L)$ имеет вид:

$$\begin{aligned} \bar{u}(x, p) = & \bar{f}(p) \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin\left(\frac{(2n-1)\pi x}{L}\right) \frac{1}{p + \frac{a\pi^2(2n-1)^2}{L^2}} + \\ & + \bar{\varphi}(p) \left(\frac{L-x}{L} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{\pi n(L-x)}{L}\right) \frac{p}{p + \frac{a\pi^2 n^2}{L^2}} \right) + \\ & + \bar{\psi}(p) \left(\frac{x}{L} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{\pi nx}{L}\right) \frac{p}{p + \frac{a\pi^2 n^2}{L^2}} \right). \end{aligned}$$

Принимая во внимание свойства функций $\varphi(t)$, $\psi(t)$ и $u(x, t)$, далее применим обратное преобразование Лапласа для $\bar{u}(x, p)$, а затем используем теорему о свертке. Получим, что решение $u(x, t)$ прямой задачи (6) - (8) имеет вид:

$$\begin{aligned} u(x, t) = & \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin\left(\frac{(2n-1)\pi x}{L}\right) e^{-\frac{(2n-1)^2 \pi^2 a}{L^2} t} \int_0^t f(\tau) e^{\frac{(2n-1)^2 \pi^2 a}{L^2} \tau} d\tau + \\ & + \left(\varphi(t) \frac{L-x}{L} + \psi(t) \frac{x}{L} \right) + \\ & + \frac{2}{\pi} \frac{d}{dt} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{\pi n(L-x)}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2} t} \int_0^t \varphi(\tau) e^{\frac{n^2 \pi^2 a}{L^2} \tau} d\tau + \\ & + \frac{2}{\pi} \frac{d}{dt} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{\pi nx}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2} t} \int_0^t \psi(\tau) e^{\frac{n^2 \pi^2 a}{L^2} \tau} d\tau. \end{aligned} \quad (10)$$

Подставляя соотношение (5) в (10), принимая во внимание то, что вместо точных значений $g(t)$ известны значения $g_\delta(t)$, после соответствующих преобразований, получаем следующее уравнение:

$$\begin{aligned} & \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin\left(\frac{(2n-1)\pi x_0}{L}\right) \int_0^t f(\tau) e^{-\frac{(2n-1)^2 \pi^2 a}{L^2} (t-\tau)} d\tau \\ & = g_\delta(t) - \left(\varphi(t) \frac{L-x_0}{L} + \psi(t) \frac{x_0}{L} \right) + \\ & + \frac{2}{\pi} \frac{d}{dt} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin\left(\frac{\pi n(L-x_0)}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2} t} \int_0^t \varphi_\delta(\tau) e^{\frac{n^2 \pi^2 a}{L^2} \tau} d\tau - \\ & + \frac{2}{\pi} \frac{d}{dt} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin\left(\frac{\pi nx_0}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2} t} \int_0^t \psi_\delta(\tau) e^{\frac{n^2 \pi^2 a}{L^2} \tau} d\tau. \end{aligned}$$

Принимая во внимание специфику сходимости рядов, стоящих в правой части последнего уравнения, применим регуляризирующий подход для раскрытия оператора дифференцирования. Суть регуляризации заключается в следующем. Обобщая результаты, представленные в работах [16, 17] и следуя идее, приведенной в [19], сначала аппроксимируем все ряды, входящие в уравнение конечными суммами, получаем:

$$\begin{aligned} & \frac{4}{\pi} \sum_{n=1}^N \frac{1}{2n-1} \sin\left(\frac{(2n-1)\pi x_0}{L}\right) \int_0^t f(\tau) e^{-\frac{(2n-1)^2 \pi^2 a}{L^2}(t-\tau)} d\tau \\ & = g_\delta(t) - \left(\varphi(t) \frac{L-x_0}{L} + \psi(t) \frac{x_0}{L} \right) + \\ & + \frac{2}{\pi} \frac{d}{dt} \sum_{n=1}^M \frac{(-1)^{n+1}}{n} \sin\left(\frac{\pi n(L-x_0)}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2} t} \int_0^t \varphi_\delta(\tau) e^{\frac{n^2 \pi^2 a}{L^2} \tau} d\tau - \\ & + \frac{2}{\pi} \frac{d}{dt} \sum_{n=1}^H \frac{(-1)^{n+1}}{n} \sin\left(\frac{\pi n x_0}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2} t} \int_0^t \psi_\delta(\tau) e^{\frac{n^2 \pi^2 a}{L^2} \tau} d\tau. \end{aligned}$$

Далее в конечных рядах раскрываем оператор дифференцирования, получаем следующее уравнение:

$$\begin{aligned} & \frac{4}{\pi} \sum_{n=1}^N \frac{1}{2n-1} \sin\left(\frac{(2n-1)\pi x_0}{L}\right) \int_0^t f(\tau) e^{-\frac{(2n-1)^2 \pi^2 a}{L^2}(t-\tau)} d\tau = \\ & = g_\delta(t) + \frac{2\pi a}{L^2} \sum_{n=1}^M n \sin\left(\frac{\pi n x_0}{L}\right) \int_0^t \varphi_\delta(\tau) e^{-\frac{n^2 \pi^2 a}{L^2}(t-\tau)} d\tau + \\ & + \frac{2\pi a}{L^2} \sum_{n=1}^H (-1)^n n \sin\left(\frac{\pi n x_0}{L}\right) \int_0^t \psi_\delta(\tau) e^{-\frac{n^2 \pi^2 a}{L^2}(t-\tau)} d\tau. \end{aligned} \quad (11)$$

Основываясь на свойствах конечных рядов, преобразовав (11), получаем

$$\begin{aligned} & \int_0^t \frac{4}{\pi} \sum_{n=1}^N \frac{1}{2n-1} \sin\left(\frac{(2n-1)\pi x_0}{L}\right) e^{-\frac{(2n-1)^2 \pi^2 a}{L^2}(t-\tau)} f(\tau) d\tau = \\ & = g_\delta(t) + \int_0^t \frac{2\pi a}{L^2} \sum_{n=1}^M n \sin\left(\frac{\pi n x_0}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2}(t-\tau)} \varphi(\tau) d\tau + \\ & + \int_0^t \frac{2\pi a}{L^2} \sum_{n=1}^H (-1)^n n \sin\left(\frac{\pi n x_0}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2}(t-\tau)} \psi(\tau) d\tau. \end{aligned}$$

Введем следующие обозначения:

$$\begin{aligned} K_N(t, \tau) &= \frac{4}{\pi} \sum_{n=1}^N \frac{1}{2n-1} \sin\left(\frac{(2n-1)\pi x_0}{L}\right) e^{-\frac{(2n-1)^2 \pi^2 a}{L^2}(t-\tau)}, \\ \Phi_{M\delta}(t) &= \int_0^t \frac{2\pi a}{L^2} \sum_{n=1}^M n \sin\left(\frac{\pi n x_0}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2}(t-\tau)} \varphi_\delta(\tau) d\tau, \\ \Psi_{H\delta}(t) &= \int_0^t \frac{2\pi a}{L^2} \sum_{n=1}^H (-1)^n n \sin\left(\frac{\pi n x_0}{L}\right) e^{-\frac{n^2 \pi^2 a}{L^2}(t-\tau)} \psi_\delta(\tau) d\tau / \end{aligned}$$

Тогда уравнение (10) примет вид:

$$\int_0^t K_N(t - \tau)f(\tau)d\tau = g_\delta(t) + \Phi_{M\delta}(t) + \Psi_{H\delta}(t). \quad (12)$$

Уравнение (11) устанавливает явную зависимость неизвестной функции источника $f(t)$ от зашумленных исходных данных. Таким образом, исходная обратная задача (1) - (5) сведена к интегральному уравнению идентификации, для которого необходимо разработать численный метод, позволяющий определять неизвестную функцию $f_\delta(t)$ по зашумленным исходным данным.

Метод идентификации функции источника и определение температурных полей во внутренних точках объекта. Определения влияния внутреннего источника на тепловое состояние объекта осуществляется в два этапа. На первом этапе происходит идентификация функции источника $f(t)$ из уравнения (12). Это уравнение является уравнением Вольтерра первого рода, относится к классу обратных задач и для численного решения этого уравнения необходимо разработать вычислительную схему, основанную на регуляризирующем алгоритме. Основная идея регуляризации заключается в том, что приближенное решение $f_\delta(t)$ уравнения (11) находят из уравнения:

$$\int_0^t K_N(t - \tau)f(\tau)d\tau + \alpha f(t) = g_\delta(t) + \Phi_{M\delta}(t) + \Psi_{H\delta}(t),$$

а настраиваемыми параметрами регуляризации являются не только α , но и количества слагаемых в ядре $K_N(t, \tau)$ и в интегралах $\Phi_{M\delta}(t)$ и $\Psi_{H\delta}(t)$. Для выбора значений параметров регуляризации α, N, M, H и оценки погрешности численного решения $f_\delta(t)$ уравнения (12) введем в рассмотрение функцию:

$$\Delta_\delta(t) = \left\| \int_0^t K_N(t - \tau)f_\delta(\tau)d\tau - (g_\delta(t) + \Phi_{M\delta}(t) + \Psi_{H\delta}(t)) \right\|$$

В предлагаемом методе значения параметров регуляризации α, N, M, H выбираются так, чтобы значения функции погрешности $\Delta_\delta(t)$ находились в допустимых пределах.

Для численной реализации метода интегралы в уравнении заменялись конечными суммами по схеме прямоугольников. Шаг дискретизации определяется временным интервалом между измерениями в граничной точке. Введение в вычислительную схему с априорно заданным шагом дискретизации стабилизирующего функционала с параметром α обеспечивает устойчивость вычислительной схемы относительно шума в исходных данных.

Отметим, что в ситуации, когда характеристики средств измерения допускают изменения временного интервала между измерениями, то для численного решения уравнения (12) возможна разработка метода, обладающего свойством саморегуляризации. Подобные методы применяются для решения различного рода задач [6].

Второй этап включает непосредственный расчет внутренних нестационарных температурных полей. Для этого переходят к решению прямой задачи (6) - (8), подставив в уравнение теплопроводности найденную функцию источника, а для численной реализации метода используют конечно-разностные схемы.

Верификация предложенного метода и оценка точности информации о внутреннем тепловом состоянии объекта осуществлялась посредством вычислительного эксперимента.

Результаты экспериментальных исследований. Основная цель эксперимента заключалась в оценке точности метода идентификации внутреннего источника и определения внутренних нестационарных температурных полей. Оценка точности проводилась на основе сравнительного анализа численных решений с тестовыми значениями, сформированными на основе имитационного моделирования. Для формирования тестовых значений внутренних температурных полей $u(x, t)$ использовали конечно-разностный аналог прямой задачи (6) -

(8). В эксперименте рассматривались различные ситуации, связанные с влиянием внутреннего теплового источника:

- Тест 1. Увеличение влияния внутреннего источника. Такая ситуация вызвана, например, процессами старения или нарушением структуры теплоизолирующих материалов в токопроводящих системах. В эксперименте эта ситуация представлена возрастающими тестовыми функциями $f(t)$.
- Тест 2. Уменьшение влияния внутреннего источника. С этим эффектом неразрывно связана первичная термообработка металлоизделий, основной целью которой является выравнивание структуры и устранение дефектов материала. В эксперименте такие ситуации представлены убывающими тестовыми функциями $f(t)$.

После идентификации функции источника вычисляли температурные значения $u_\delta(x, t)$, подставляя найденную функцию источника $f_\delta(t)$ в конечно-разностный аналог прямой задачи (6) - (8). Для оценки отклонений вычисленной функции источника $f_\delta(t)$ от тестовых значений $f(t)$ используем функцию $\Delta(t)$ и величины Δ_f и ϑ_f :

$$\Delta(t) = |f_\delta(t) - f(t)|, \quad \Delta_f = \max_{t \in [0, T]} \Delta(t), \quad \vartheta_f = \frac{\Delta_f}{\max_{t \in [0, T]} |f(t)|}$$

Далее оценивали отклонение функции $u_\delta(x, t)$ от тестовых значений $u(x, t)$ с помощью величин Δ_u и ϑ_u , определяемых формулами:

$$\Delta_u = \max_{(x,t) \in [0, L] \times [0, T]} |u_\delta(x, t) - u(x, t)|, \quad \vartheta_u = \frac{\Delta_u}{\max_{(x,t) \in [0, L] \times [0, T]} |u(x, t)|}$$

Результаты эксперимента при уровне шума $\delta = 0,05$ проиллюстрированы рис.1 - рис.6. Результаты эксперимента, связанного с увеличением влияния внутреннего теплового источника, представлены на рис.1 - рис.3. Результаты, связанные с уменьшением его влияния, приведены на рис.4 - рис.6.

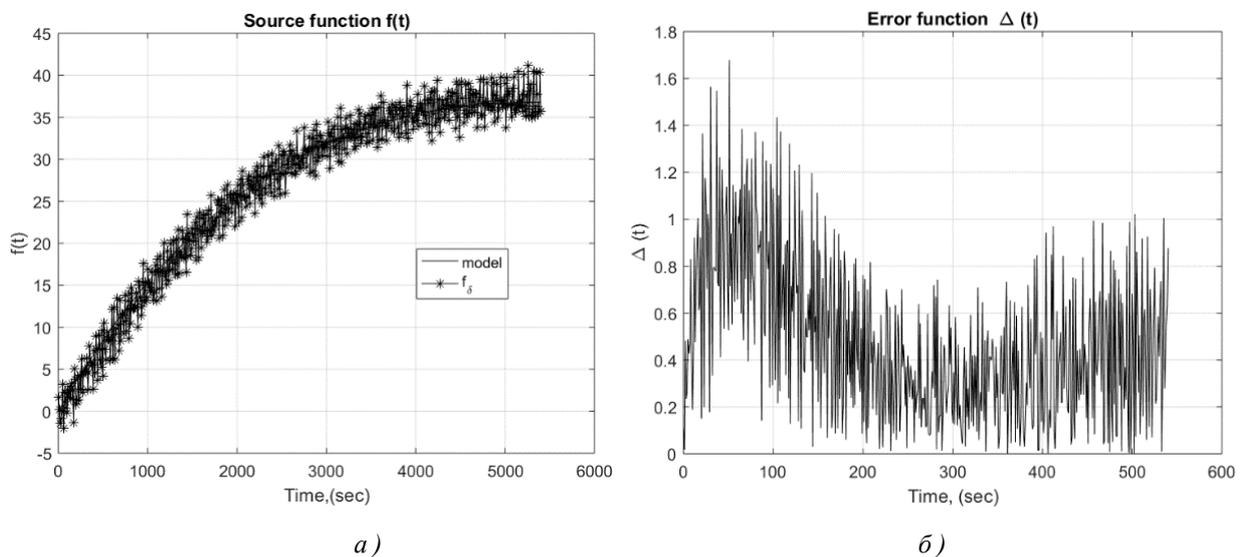


Рис. 1. Результаты идентификации функции источника (Тест 1). (а) Графики функции источника: f_δ —функция источника $f_\delta(t)$; model - тестовая функция $f(t)$; (б). График функции $\Delta(t)$

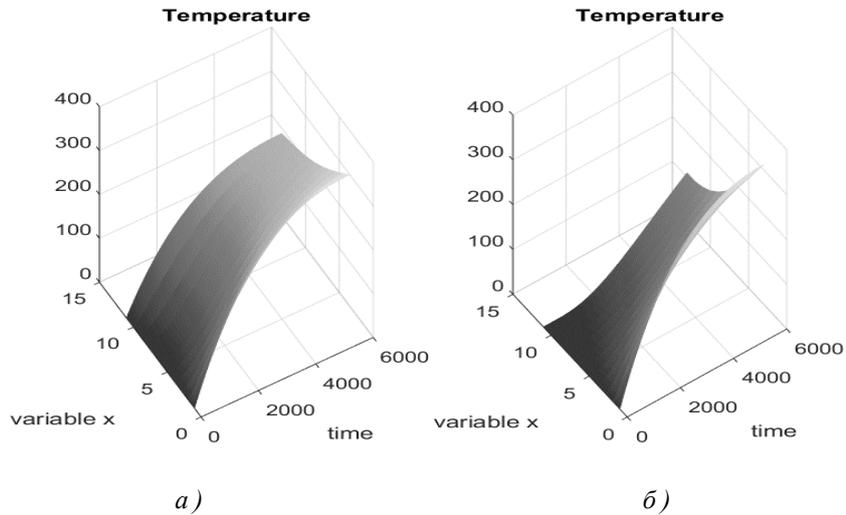


Рис. 2. Температурные поля в объекте (Тест 1).

(а) – температурные поля $u_0(x, t)$, формируемые в объекте, не содержащем внутреннего теплового источника; (б) – температурные поля $u_\delta(x, t)$, сформированные под влиянием внутреннего теплового источника

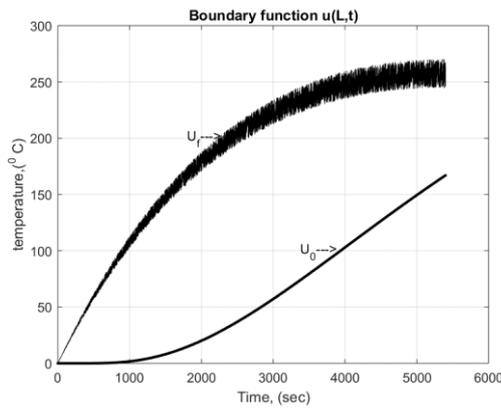


Рис. 3. Графики температур в точке $x = L$ (Тест1). U_0 – температурное поле, получаемое при отсутствии внутреннего теплового источника, U_f – температурное поле, формируемое под влиянием внутреннего теплового источника

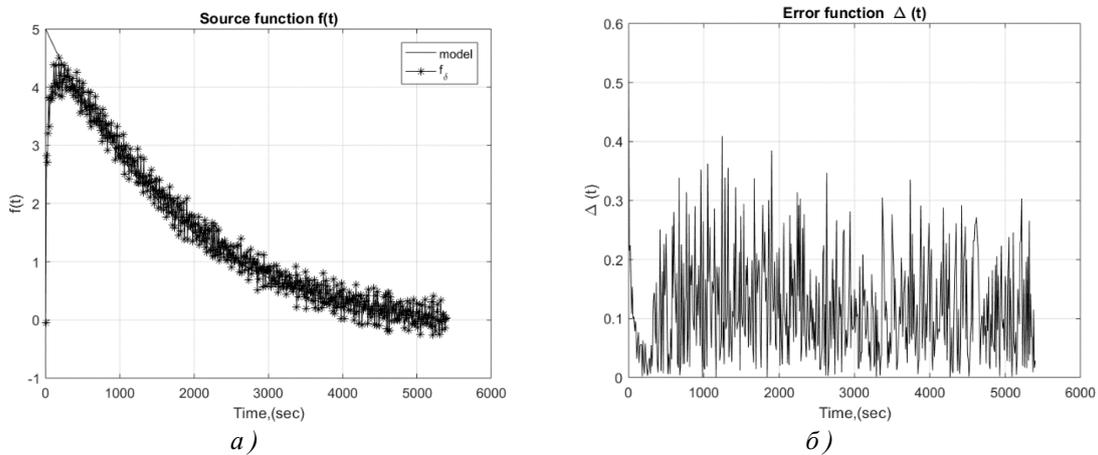


Рис. 4. Результаты идентификации функции источника (Тест 2). (а) Графики функции источника: f_δ – функция источника $f_\delta(t)$; model - тестовая функция $f(t)$; (б). График функции $\Delta(t)$

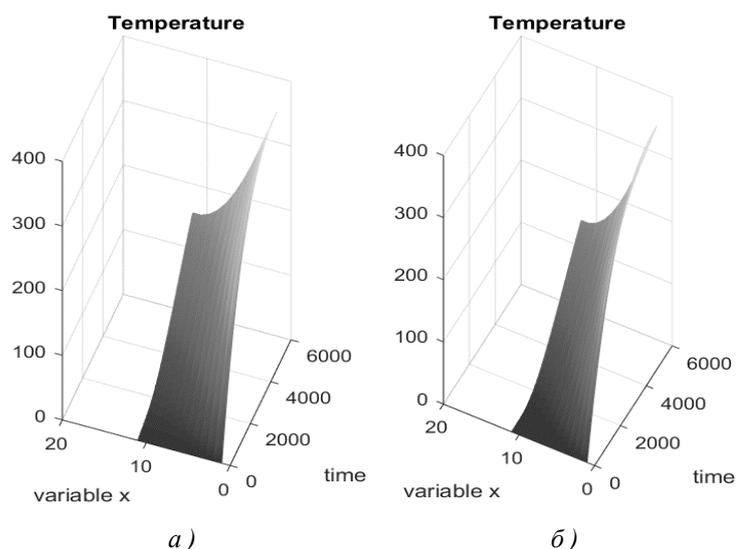


Рис. 5. Температурные поля в объекте (Тест 2). (а) – температурное поле $u_0(x, t)$, формируемые в объекте, не содержащем теплового источника; (б) – температурное поле $u_\delta(x, t)$ сформированное под влиянием внутреннего теплового источника

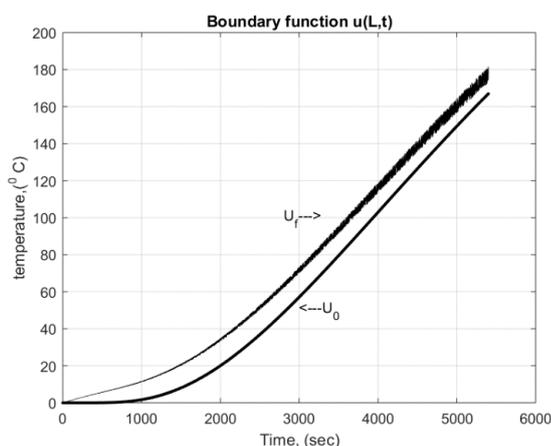


Рис. 6. Графики температур в точке $x = L$ (Тест 2). U_0 - температурное поле, получаемое при отсутствии внутреннего теплового источника, U_f - температурное поле, формируемое под влиянием внутреннего теплового источника

Результаты сравнительного анализа при различных уровнях шума исходных данных приведены в таблице 1.

Таблица 1. Оценки погрешностей тепловых функций

Тестовые функции	Уровень шума δ	Δ_f	ϑ_f	Δ_u	ϑ_u
Тест 1	0,03	1,5045	0,0231	3,4047	0,0258
	0,05	1,6742	0,0242	3,4499	0,0271
	0,1	1,9359	0,0251	3,4863	0,0281
Тест 2	0,03	0,8130	0,0274	2,3180	0,0434
	0,05	0,8221	0,0289	2,6421	0,0489
	0,1	0,8508	0,0204	2,7308	0,0514

Результаты эксперимента свидетельствуют о принципиальной возможности идентификации внутреннего источника с помощью предложенного подхода в условиях проведения динамических измерений, а также об устойчивости и достаточной точности метода идентификации источника и определения нестационарных внутренних температурных полей относительно шума исходных данных.

Заключение. В статье предложены интегральная модель и метод идентификации внутреннего теплового источника по результатам граничных измерений. Интегральная модель, полученная на основе преобразований Лапласа и регуляризирующих подходов, устанавливает прямую зависимость неизвестной функции источника от исходных данных, формируемых из результатов граничных температурных измерений. В работе предложен численный метод решения интегрального уравнения, основанный на регуляризирующем подходе, обеспечивающем уменьшение влияния шумов на устойчивость и точность результатов идентификации внутреннего теплового источника, а также алгоритм определения теплового состояния объекта. Надёжность метода идентификации и достоверность получаемой информации о тепловом состоянии объекта подтверждается результатами эксперимента.

Благодарности. Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (государственное задание №FENU-2021-0014).

Список источников

1. Efendiev Y., Pun S.M., Vabishchevich P.N. Temporal splitting algorithms for non-stationary multiscale problems. *Journal of Computational Physics*, 2021, vol. 439, iss. 1, 110375.
2. Vasyliiev V., Vasilyeva M., An accurate approximation of the two-phase stefan problem with coefficient smoothing. *Mathematics*, 2020, vol. 8, iss. 11, 19224.
3. Lukyanenko D.V., Borzunov A.A., and Shishlenin M.A. Solving coefficient inverse problems for nonlinear singularly perturbed equations of the reaction-diffusion-advection type with data on the position of a reaction front. *Communications in nonlinear science and numerical simulation*, 2020, vol. 20, iss. 4, pp. 727-737.
4. Korotkii A.I. and Starodubtseva Y.V. Direct and inverse boundary value problems for models of stationary reaction–convection–diffusion. *Proceedings of the Steklov Institute of Mathematics*, 2015, vol. 291, pp. 96-112.
5. Дилигенская А.Н. Решение ретроспективной обратной задачи теплопроводности на основе параметрической оптимизации / А.Н. Дилигенская // *Теплофизика высоких температур*, 2018. – Т. 56. – № 3. – С. 399-406.
6. Procudina L.A., Yaparov D. Numerical modeling of the mass of the flowing liquid at transverse oscillations of the straight tube. *AIP conference proceedings application of mathematics in technical and natural sciences: 11th international conference for promoting the application of mathematics in technical and natural sciences*, 2019, 2164 – 120011.
7. Franco D., Infante G., Minhos F.M. Nonlocal boundary value problems. *Boundary value problems*, 2012.
8. Yaparova N. Mathematical modelling and method for solving a parametric identification problem for self-test of measuring devices. *Inverse problems in science and engineering*, 2016, vol. 24, iss. 1, pp. 77-91.
9. Bazhlekova E., Bazhlekov I. Identification of a space-dependent source term in a nonlocal problem for the general time-fractional diffusion equation. *Journal of computational and applied mathematics*, 2021, vol. 384, 113213.
10. Slimani S., Medarhi I., Najib K., Zine A. Identification of the source function for a seawater intrusion problem in unconfined aquifer. *Numerical algorithms*, 2020, vol. 84, iss. 4, pp. 1565-1587.
11. Rostamian M., Shahrezaee A. A meshless method for solving 1D time-dependent heat source problem. *Inverse problems in science and engineering*, 2018, vol. 26, iss. 1, pp. 51-81.
12. Ismailov M.I., Kanca F., Lesnic D. Determination of a time-dependent heat source under nonlocal boundary and integral overdetermination conditions. *Applied mathematics and computation*, 2011, vol. 218, pp. 4138-4146.
13. Zhou G., Wu B. Identifying a time-dependent heat source with nonlocal boundary and overdetermination conditions by the variational iteration method. *International journal of numerical methods for heat fluid flow*, 2014, vol. 24, iss. 7, pp. 1545-1552.
14. Farcas A., Lesnic D. The boundary-element method for the determination of a heat source dependent on one variable. *Journal of engineering mathematics*, 2006, vol. 54, pp. 375-388.
15. Прилепко А.И. Корректность обратной задачи об источнике для параболических систем / А.И. Прилепко, Д.С. Ткаченко // *Дифференциальные уравнения*, 2004. – Т. 40. – № 11. – С. 1540-1547.
16. Yaparova N. Numerical methods for solving a boundary-value inverse heat conduction problem. *Inverse Problems in science and engineering*, 2014, vol. 22, no. 5, pp. 832-847.
17. Япарова Н.М. Метод решения обратной задачи идентификации функции источника с использованием преобразования Лапласа / Н.М. Япарова // *Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика*, 2016. – Т. 5. – № 3. – С. 20-35.

18. Лаврентьев М.А. Методы теории функции комплексного переменного / М.А. Лаврентьев, Б.В. Шабат. – М.: Наука, 1973. – 738 с.
19. Лаврентьев М.М. Некорректные задачи математической физики и анализа / М.М. Лаврентьев, В.Г. Романов, С.П. Шишатский. – М.: Наука, 1980. – 286 с.

Япарова Наталья Михайловна. Доктор технических наук, доцент, заведующий кафедрой математического обеспечения информационных технологий Южно-Уральского государственного университета (НИУ), Основные научные интересы: методы обработки информации, разработка и исследование методов и алгоритмов решения задач тепло и массопереноса с неполными, зашумленными исходными данными, обратные задачи. AuthorID: 15799, SPIN: 6353-4685, ORCID: 0000-0003-1821-1159, iaparovam@susu.ru, 454080, Россия, г. Челябинск, пр. Ленина, 76.

Капелюшин Юрий Евгеньевич. PhD, доцент, заведующий лабораторией научно-образовательного центра Южно-Уральского государственного университета (НИУ). Основные научные интересы: процессы тепло и массопереноса, создание новых производственных технологий и материалов. AuthorID: 955090, SPIN: 2543-2742, ORCID: 0000-0002-3352-1780, kapelyushinye@susu.ru, 454080, Россия, г. Челябинск, пр. Ленина, 76.

UDC 519.6+517.9

DOI:10.25729/ESI.2023.32.4.008

Identification of an internal heat source and determining the thermal state of an object by dynamic boundary measurements

Natalia M. Yaparova, Yury E. Kapelyushin

South Ural State University (NRU),

Russia, Chelyabinsk, iaparovam@susu.ru

Abstract. The article considers the problem of identifying an internal heat source and assessing its influence on changes in the temperature of a controlled object. The problem of identifying a heat source arises during thermal control of power transmission, during heat treatment, and during non-destructive testing of buildings, structures and materials. The mathematical model of heat transfer inside an object is represented by a heat conduction equation with an unknown source function, initial conditions, and boundary conditions formed on the basis of the results of noisy temperature measurements obtained near the surface of the object. The article proposes an approach to identifying an internal heat source based on the transition from an inverse problem to an integral equation, a numerical method for solving it, as well as an algorithm for calculating non-stationary internal heat fields that takes into account the influence of a heat source. The stability of the identification method with respect to the error of the initial data is ensured by the choice of regularization parameters. The proposed approach and methods, in contrast to the existing ones, make it possible to establish an explicit dependence of the desired function of the internal heat source on boundary measurements in a situation where the temperature fields near the surface of the object change with time. The article presents error estimates for numerical solutions found as a result of a comparative analysis with test values. The results of the experiment indicate that the proposed methods reduce the negative impact of noise on the accuracy of data processing and allow determining the internal thermal state of an object from indirect measurements with a sufficient level of accuracy and can serve as a basis for determining the influence of an internal heat source on the formation of internal non-stationary temperature fields.

Keywords: heat transfer, technical systems, information processing method, inverse problem, integral equation, numerical method, regularization

Acknowledgements: The research was carried out under State Assignment Project (no. FWEU-2021-0014) of the Fundamental Research Program of Russian Federation 2021-2030 and funded by RFBR.

References

1. Efendiev Y., Pun S.M., Vabishchevich P.N. Temporal splitting algorithms for non-stationary multiscale problems. *Journal of Computational Physics*, 2021, vol. 439, iss. 1, 110375.
2. Vasylyev V., Vasilyeva M., An accurate approximation of the two-phase stefan problem with coefficient smoothing. *Mathematics*, 2020, vol. 8, iss. 11, 19224.

3. Lukyanenko D.V., Borzunov A.A., and Shishlenin M.A. Solving coefficient inverse problems for nonlinear singularly perturbed equations of the reaction-diffusion-advection type with data on the position of a reaction front. *Communications in nonlinear science and numerical simulation*, 2020, vol. 20, iss. 4, pp. 727-737.
4. Korotkii A.I. and Starodubtseva Y.V. Direct and inverse boundary value problems for models of stationary reaction-convection-diffusion. *Proceedings of the Steklov Institute of Mathematics*, 2015, vol. 291, pp. 96-112.
5. Diligenskaya A.N. Reshenie retrospektivnoj obratnoj zadachi teploprovodnosti na osnove parametriceskoj optimizacii [Solution of the retrospective inverse heat conduction problem with parametric optimization]. *Teplofizika vysokikh temperature [Thermophysics of high temperatures]*, 2018, vol. 56, no 3, pp. 382-388.
6. Procudina L.A., Yaparov D. Numerical modeling of the mass of the flowing liquid at transverse oscillations of the straight tube. *AIP conference proceedings application of mathematics in technical and natural sciences: 11th international conference for promoting the application of mathematics in technical and natural sciences*, 2019, 2164 – 120011.
7. Franco D., Infante G., Minhos F.M. Nonlocal boundary value problems. *Boundary value problems*, 2012.
8. Yaparova N. Mathematical modelling and method for solving a parametric identification problem for self-test of measuring devices. *Inverse problems in science and engineering*, 2016, vol. 24, iss. 1, pp. 77-91.
9. Bazhlekova E., Bazhlekov I. Identification of a space-dependent source term in a nonlocal problem for the general time-fractional diffusion equation. *Journal of computational and applied mathematics*, 2021, vol. 384, 113213.
10. Slimani S., Medarhri I., Najib K., Zine A. Identification of the source function for a seawater intrusion problem in unconfined aquifer. *Numerical algorithms*, 2020, vol. 84, iss. 4, pp. 1565-1587.
11. Rostamian M., Shahrezaee A. A meshless method for solving 1D time-dependent heat source problem. *Inverse problems in science and engineering*, 2018, vol. 26, iss. 1, pp. 51-81.
12. Ismailov M.I., Kanca F., Lesnic D. Determination of a time-dependent heat source under nonlocal boundary and integral overdetermination conditions. *Applied mathematics and computation*, 2011, vol. 218, pp. 4138-4146.
13. Zhou G., Wu B. Identifying a time-dependent heat source with nonlocal boundary and overdetermination conditions by the variational iteration method. *International journal of numerical methods for heat fluid flow*, 2014, vol. 24, iss. 7, pp. 1545-1552.
14. Farcas A., Lesnic D. The boundary-element method for the determination of a heat source dependent on one variable. *Journal of engineering mathematics*, 2006, vol. 54, pp. 375-388.
15. Prilepko A.I., Tkachenko D.S. Korrektnost' obratnoj zadachi ob istochnike dlya parabolicheskikh sistem [Well-posedness of the inverse source problem for parabolic systems]. *Differentsial'nyye uravneniya [Differential Equations]*, 2004, vol.40, iss. 11, pp. 1619-1626
16. Yaparova N. Numerical methods for solving a boundary-value inverse heat conduction problem. *Inverse problems in science and engineering*, 2014, vol. 22, no. 5, pp. 832-847.
17. Yaparova N.M. Metod resheniya obratnoj zadachi identifikacii funkicii istochnika s ispol'zovaniem preobrazovaniya Laplasya [Method for solving an inverse term source problem based on the Laplace transform]. *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya: Vychislitel'naya matematika i informatika [Bulletin of the South Ural state university. Series: Computational mathematics and software engineering]*, 2016, vol. 5, iss. 3, pp. 20-35.
18. Lavrent'ev M.A., Shabat B.V. *Metody teorii funkicii kompleksnogo peremennogo [Methods of the theory of functions of a complex variable]*. Moscow, Nauka [Science], 1973, 738 p.
19. Lavrent'ev M.M., Romanov V. G., Shishatskii S. P. *Nekorrektnye zadachi matematicheskoy fiziki i analiza [Ill-posed problems of mathematical physics and analysis]*. Moscow, Nauka [Science], 1980. 286 p.

Yaparova Natalia Mikhailovna. Doctor of Science, Associate Professor, Head of the Department of Mathematical Foundations of Information Technologies, South Ural State University (NRU). Main scientific areas of interest: Data processing, methods and algorithms for solving the heat and mass transfer problems with incomplete and noisy initial data, inverse problems. AuthorID: 15799, SPIN: 6353-4685, ORCID: 0000-0003-1821-1159, iaparovannm@susu.ru, 454080, 76, Lenin prospekt, Chelyabinsk, Russia.

Kapelyushin Yury Evgenyevich. PhD, Associate Professor, Head of the laboratory of the scientific and educational center, South Ural State University (NRU). Main scientific areas of interest: heat and mass transfer processes, creation of new production technologies and materials. AuthorID: 955090, SPIN: 2543-2742, ORCID: 0000-0002-3352-1780, kapelyushinye@susu.ru, 454080, 76, Lenin prospekt, Chelyabinsk, Russia.

Статья поступила в редакцию 12.04.2023; одобрена после рецензирования 28.11.2023; принята к публикации 16.12.2023.

The article was submitted 04/12/2023; approved after reviewing 11/28/2023; accepted for publication 12/16/2023.

Исследование динамики реакторной установки ВВЭР-СКД при различных изменениях параметров питательной воды

Суджян Артавазд Манукович

Национальный исследовательский центр «Курчатовский институт»,

Россия, Москва, artavazd1994@gmail.com

Аннотация. В статье представлены результаты расчетных работ по исследованию динамики реакторной установки ВВЭР-СКД, выполненных с использованием возможностей программного кода RELAP5/MOD3.3. Характерной особенностью данного реактора является довольно заметное изменение плотности теплоносителя в активной зоне. В сочетании с нейтронно-физическими обратными связями создаются риски возникновения теплогидравлической и нейтронно-теплогидравлической неустойчивости, в особенности при переходных процессах. В связи с этим, необходимо иметь возможности прогнозировать возникновение неустойчивостей, оценивать границы устойчивости системы. Для этих целей была разработана расчетная модель реакторной установки. Выбор для данной работы кода RELAP5 обусловлен широким опытом его использования при обосновании безопасности существующих водоохлаждаемых реакторов. Для получения возможности расчетов ВВЭР-СКД по коду RELAP5 свойства воды были расширены и детализированы в области сверхкритических давлений и температур. Для разработанной модели реакторной установки ВВЭР-СКД были выполнены расчеты условного пуска реактора с последующим выходом на номинальный режим работы. Также было рассмотрено влияние отклонений температуры и расхода питательной воды на динамику реакторной установки. По результатам расчетов сделаны выводы об устойчивости работы реакторной установки в номинальном режиме.

Ключевые слова: реакторная установка, ВВЭР-СКД, динамика, устойчивость, питательная вода, обратные связи

Цитирование: Суджян А.М. Исследование динамики реакторной установки ВВЭР-СКД при различных изменениях параметров питательной воды / А.М. Суджян // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 104-116. – DOI:10.25729/ESI.2023.32.4.009.

Введение. Ядерные энергетические установки (ЯЭУ), охлаждаемые водой при сверхкритических параметрах, входят в число шести типов ядерных реакторов, представляющих четвертое поколение [1]. В связи с этим, в ряде стран проводится разработка своей концепции реактора данного типа. В России в настоящее время выполняются НИОКР по разработке технологий реактора ВВЭР-СКД. Рассматриваемая концепция представляет собой двухконтурную ЯЭУ с быстро-резонансным спектром нейтронов, с водой при сверхкритических параметрах как в первом, так и во втором контуре [2, 3]. Из-за довольно сильного уменьшения плотности теплоносителя в активной зоне реактора, а также положительной обратной связи от плотностного коэффициента реактивности, при определенных условиях в нем может возникнуть колебательная и экспоненциальная неустойчивость [4, 5]. Для решения данных проблем на данном этапе проводятся расчетные работы по исследованию динамики реакторной установки (РУ) ВВЭР-СКД, для чего моделируются различные переходные процессы, оценивается влияние обратных связей на устойчивость работы реактора, определяются границы областей неустойчивости.

Область псевдофазового перехода сверхкритических флюидов характеризуется всплеском теплоемкости, сжимаемости, а также существенным изменением плотности и кинематических свойств. Состояние воды при низкой плотности называется псевдопаровым, а при высокой – псевдожидким [6]. Для снижения вероятности возникновения неустойчивостей, предполагаемые рабочие параметры первого контура ВВЭР-СКД находятся в псевдопаровой области, а влияние плотностного эффекта реактивности предполагается скомпенсировать доплеровским.

Одной из задач исследования динамики реактора является оценка отклика системы на изменение параметров питательной воды. Снижение или увеличение температуры и массового

расхода питательной воды приводят к изменениям условий теплообмена между первым и вторым контуром. Происходит изменение всех температур в реакторной установке, а вместе с тем и изменение мощности через обратные связи.

1. Существующий опыт. Несмотря на то, что аналогичных работ для ВВЭР-СКД не существует, в качестве примера можно рассматривать выполненные исследования для других концепций реакторов со сверхкритическими параметрами теплоносителя.

В работе [7] приведен анализ поведения температур оболочки тепловыделяющих элементов (ТВЭЛ) концепции реактора с тепловым спектром нейтронов US SCWR при остановке циркуляции питательной воды, снижении температуры питательной воды на 30 °С (с 280 °С до 250 °С), а также при полном разрыве линии подачи питательной воды. Расчеты выполнены с помощью модифицированного программного кода RELAP5-3D. В первом случае анализ показал, что отключение насоса питательной воды не критично, поскольку происходит быстрое срабатывание клапанов сброса пара и нивелируется избыточное давление в системе, по аналогии с кипящим реактором BWR (boiling water reactor). Однако вклад реактивности не так высок, как в BWR, поскольку большая часть замедления в активной зоне SCWR (supercritical water reactor) достигается за счет жидкого теплоносителя (водяных стержней), на который не влияет повышение давления. В результате топливо не перегревается. При снижении температуры питательной воды на 30 С топливо и оболочка ТВЭЛ также не перегреваются за счет более низкой температуры теплоносителя и доплеровского эффекта реактивности. В случае полного разрыва линии подачи питательной воды (в отсутствие систем аварийного охлаждения) температура оболочки ТВЭЛ достигает аварийного предела в 1205 °С примерно за 25 секунд. Следовательно, необходимо определить систему впрыска высокого давления (активную или пассивную) для первоначального смягчения последствий, а также и систему низкого давления (активную или пассивную, со сбросом давления и гравитационным впрыском).

Анализ типичных возмущений концепции быстрого реактора SUPER FR представлен в работе [8]. Рассмотрены случаи импульсного увеличения расхода питательной воды на 5 % и поэтапного снижения температуры питательной воды на 10 °С. Анализ выполнялся с помощью системного кода SPRAT-F. В этой работе упор делается на модификацию системы контроля подачи питательной воды, конкретика по температурам не приведена. В первом случае расход питательной воды ступенчато уменьшается на 5 %, а затем восстанавливается регуляторами питательной воды. С помощью модифицированной системы управления восстановление расхода происходит быстрее и, следовательно, изменение температуры основного пара меньше, чем в случае с исходной системой управления. Реактивность, введенная за счет обратной связи по плотности теплоносителя в Super FR незначительна. Во втором случае температура питательной воды ступенчато снижается на 10 С и не восстанавливается. Сначала температура основного пара повышается, поскольку уменьшение объемного расхода от насосов питательной воды приводит к уменьшению расхода на входе в активную зону. Мощность реактора снижается из-за обратной связи по плотности теплоносителя. Затем температура основного пара снижается, а мощность реактора увеличивается после того, как холодная питательная вода начинает поступать в активную зону. При использовании систем управления все параметры возвращаются к исходным значениям в течение 2 минут.

Для концепции реактора CSR1000 с тепловым спектром нейтронов выполнен расчетный анализ безопасности переходных процессов, связанных с изменением параметров питательной воды, анализ представлен в работе [9]. Расчеты были выполнены с помощью специально разработанного кода SCAC-CSR1000. Рассмотрена ситуация снижения температуры питательной воды на 35 С в результате отключения одной ступени нагрева. В течение первых

5 секунд после отключения наблюдается кратковременное снижение расхода питательной воды, температура оболочки ТВЭЛ при этом быстро растет. После этого срабатывают системы безопасности, что позволяет не допустить рост максимальной температуры оболочки выше 700 С, при критерии безопасности 1260 С.

Можно сделать вывод о том, что на сегодняшний день работ подобного рода крайне мало, а большая часть из них выполнена для одноконтурных реакторов с тепловым спектром нейтронов, что ограничивает количество полезной информации, которую можно использовать при разработке концепции ВВЭР-СКД, в силу его двухконтурности и отличающихся эффектов реактивности. Следовательно, требуется больше исследований, учитывающих эти особенности реакторной установки (РУ).

2. Описание расчетной модели. Для исследования динамики РУ ВВЭР-СКД были использованы возможности программного кода RELAP5/MOD3.3 [10]. Это одномерный теплогидравлический код, часто применяемый при решении задач по исследованию переходных процессов в реакторах с легководным теплоносителем [11]. В RELAP5 используются теплофизические свойства воды из базы NIST [12], с охватом сверхкритической области, что позволяет использовать данный программный код в качестве расчетного инструмента для РУ ВВЭР-СКД.

Разработанная для исследования динамики модель реактора содержит основные элементы первого контура и часть второго контура, представленного парогенератором. Первый контур, в соответствии с компоновкой РУ, состоит из реактора, четырех насосов, двух парогенераторов и двух буферных емкостей. Соединение элементов выполнено с помощью трубопровода холодного и горячего участков. Нодализация схемы модели РУ ВВЭР-СКД представлена на рисунке 1.

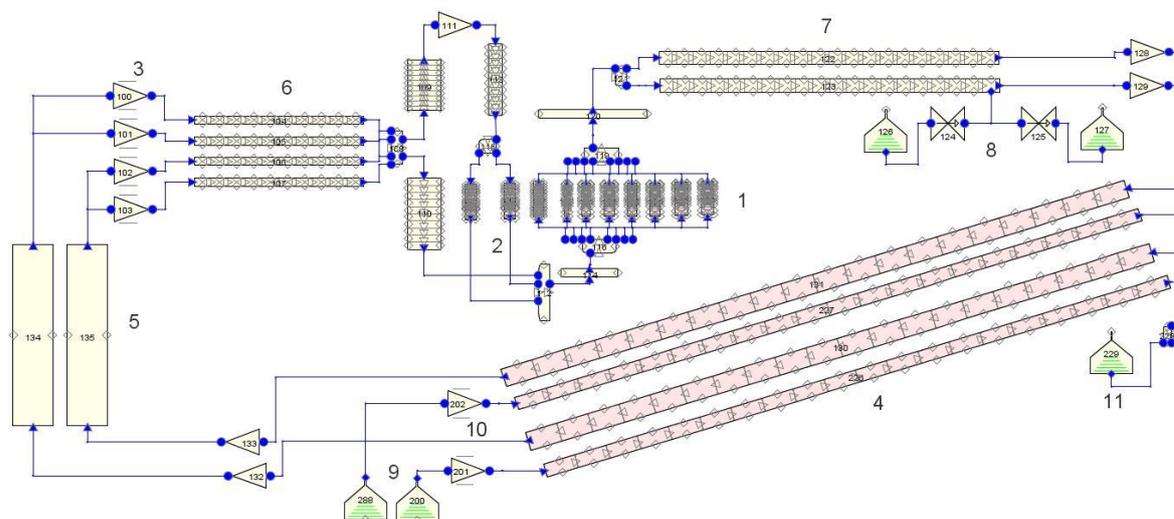


Рис. 1. Нодализация схема РУ ВВЭР-СКД в RELAP5

1 – активная зона, 2 – БЗВ, 3 – ГЦН, 4 – парогенератор, 5 – буферная емкость, 6 – холодный трубопровод, 7 – горячий трубопровод, 8 – система контроля давления, 9 – вход в парогенератор по второму контуру, 10 – насос питательной воды, 11 – выход из парогенератора по второму контуру

Управление давлением в системе осуществляется на участке горячего трубопровода с помощью регулирующих клапанов и емкостей с водой (для накачки системы) и атмосферным воздухом (для сброса избыточного давления). Клапан, предназначенный для накачки давления, начинает открываться при снижении давления до 27,47 МПа и открывается полностью при достижении 27,4 МПа. Клапан, предназначенный для сброса избыточного

давления, начинает открываться при давлении 27,7 МПа и находится в полностью открытом состоянии при достижении 28,5 МПа. В качестве граничных условий выбраны температура и расход питательной воды на входе в парогенератор (320 °С и 370 кг/с соответственно) и ее давление на выходе из парогенератора (28 МПа). В первом контуре задается расход ГЦН (435 кг/с). В качестве начальных значений давления и температуры в первом контуре используются 27,5 МПа и 405 °С соответственно. Первый и второй контур связаны посредством теплообмена между водой первого и второго контура в парогенераторах. Начальная мощность – нулевая.

3. Исследуемые переходные процессы. Для того, чтобы оценить влияние изменения параметров питательной воды на нейтронно-теплогидравлическую устойчивость ВВЭР-СКД, были выполнены расчеты, порядок которых выглядит следующим образом.

На начальном этапе выполнен условный пуск реактора, при котором устанавливается номинальный режим работы. Граничные и начальные условия соответствуют представленным в разделе «3». Мощность реактора повышается до номинальной, равной 1250 МВт, путем линейного ввода реактивности 1,362 \$ стержнями управления в течение первых 100 секунд расчета и регулировки входной температуры питательной воды. После этого система приходит к устойчивому стационарному состоянию, соответствующему номинальному режиму работы реактора. Далее это состояние используется как стартовое для расчета переходных процессов, рассматриваемые возмущения вносятся в момент времени 2500 с.

В качестве первых двух возмущений рассматривались снижение температуры питательной воды на входе в парогенератор на 40 °С с 320 °С до 280 °С и ее повышение на 30 °С с 320 °С до 350 °С. Еще два возмущения были внесены по массовому расходу питательной воды на входе в парогенератор: повышение расхода на 10 % с 370 кг/с до 405 кг/с и его снижение на 12,5 % с 370 кг/с до 324 кг/с.

Предполагалось, что время внесения каждого возмущения составляет 50 с. Фиксировались соответствующие изменения мощности, реактивностей, а также температур теплоносителя, топлива и оболочек ТВЭЛ.

4. Результаты расчета.

4.1. Пуск реактора. Поскольку при указанных ранее начальных и граничных условиях в процессе старта реактора (линейного введения реактивности за счет регулирующих стержней и набора мощности, рисунок 2а) начинается интенсивный теплообмен между контурами, в результате чего наблюдается резкое падение температуры и давления в первом контуре, для данного исследования было принято решение выполнять условный пуск реактора при высокой температуре питательной воды на входе в парогенератор по второму контуру, соответствующей стартовому состоянию первого контура (405 °С), и постепенно снижать ее до требуемой величины 320 °С по мере выхода на номинальную мощность. Таким образом, в течение первых 700 секунд параметры системы устанавливаются в соответствии с высокой температурой питательной воды на входе в парогенератор (405 °С), затем, каждые 200 секунд происходит последовательное снижение температуры питательной воды до 395 °С (псевдокритическая точка), 374 °С (критическая температура), 373 °С (докритическая температура), и в момент времени 1500 с температура питательной воды равна 320 °С. До момента времени 2000 с параметры системы устанавливаются, достигая стационарного режима. Описанный процесс снижения температуры питательной воды на входе в парогенератор для достижения номинального режима работы проиллюстрирован на рисунке 2б.

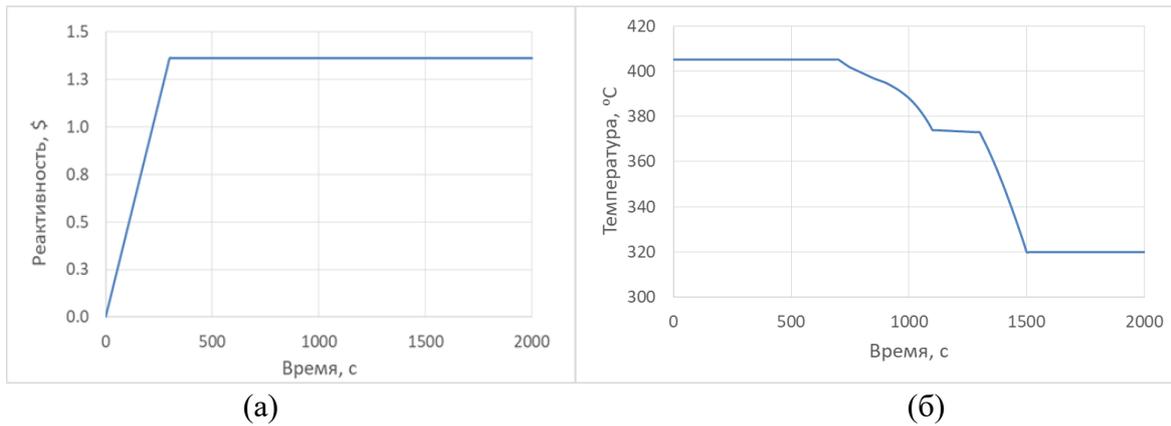


Рис. 2. Введение реактивности стержнями СУЗ (а) и изменение температуры питательной воды на входе в парогенератор (б) при пуске реактора

На рисунке 3 представлено установление в системе номинальной мощности 1250 МВт и соответствующих ей значений максимальной и минимальной температур топлива.

Максимальная температура топлива при номинальном режиме работы равна 1213 °С, минимальная – 624 °С. При этом, зависимости максимальной и минимальной температур оболочки ТВЭЛ, а также температуры теплоносителя на входе в реактор и выходе из него от времени, представлены на рисунке 4.

Полученные значения для максимальной и минимальной температур оболочки твэл в номинальном режиме составили 604 °С и 409 °С соответственно, температура теплоносителя на входе в реактор при этом равна 405 °С, а на выходе – 520 °С, что согласуется с характеристиками концепции ВВЭР-СКД.

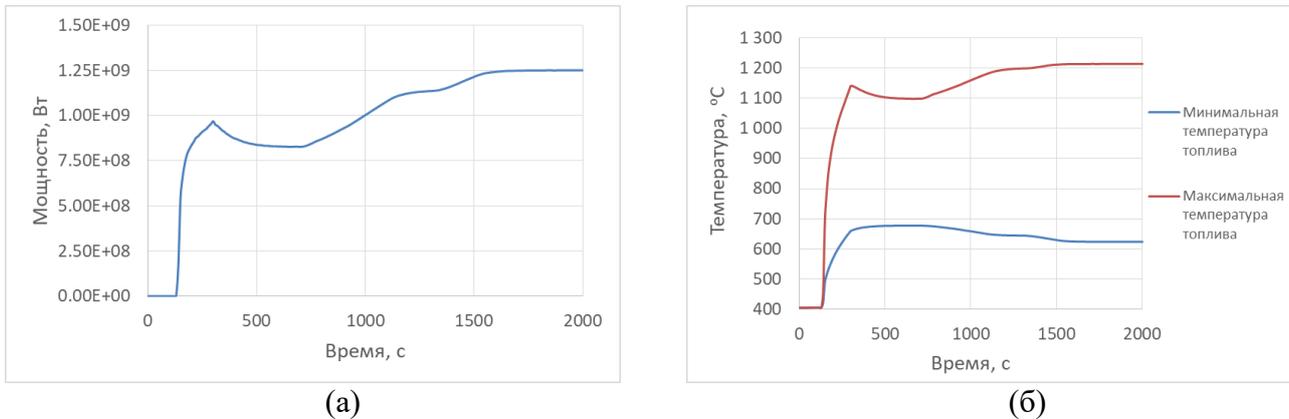


Рис. 3. Установление номинального значения мощности (а) и соответствующих максимальной и минимальной температур топлива (б)

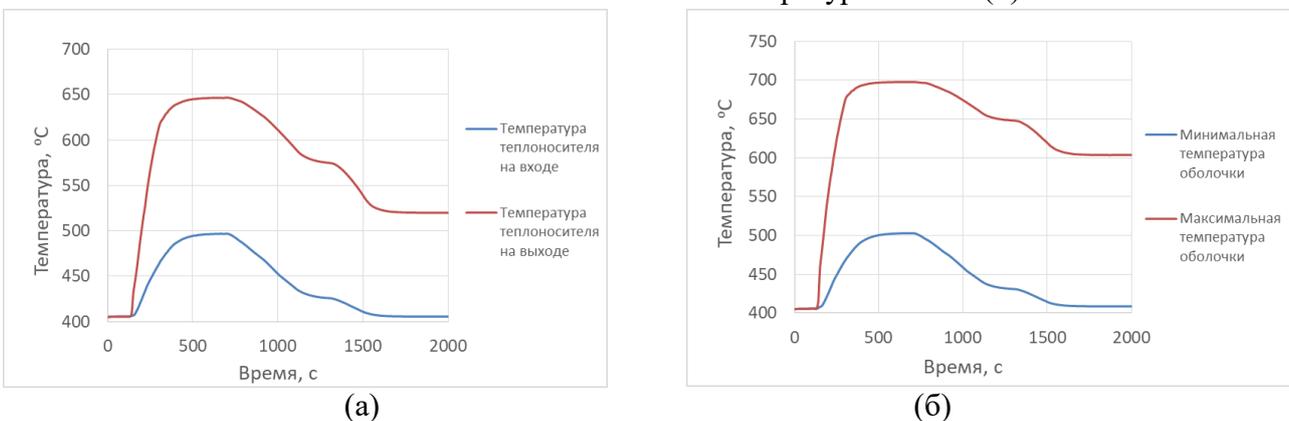


Рис. 4. Изменение температуры теплоносителя на входе в реактор и выходе из него (а) и максимальной и минимальной температуры оболочки твэл (б)

Изменения всех учитываемых в модели эффектов реактивности: введенной регулируемыми стержнями, вызванной изменением температуры топлива, плотности теплоносителя, а также суммарной, в зависимости от времени, представлены на рисунке 5.

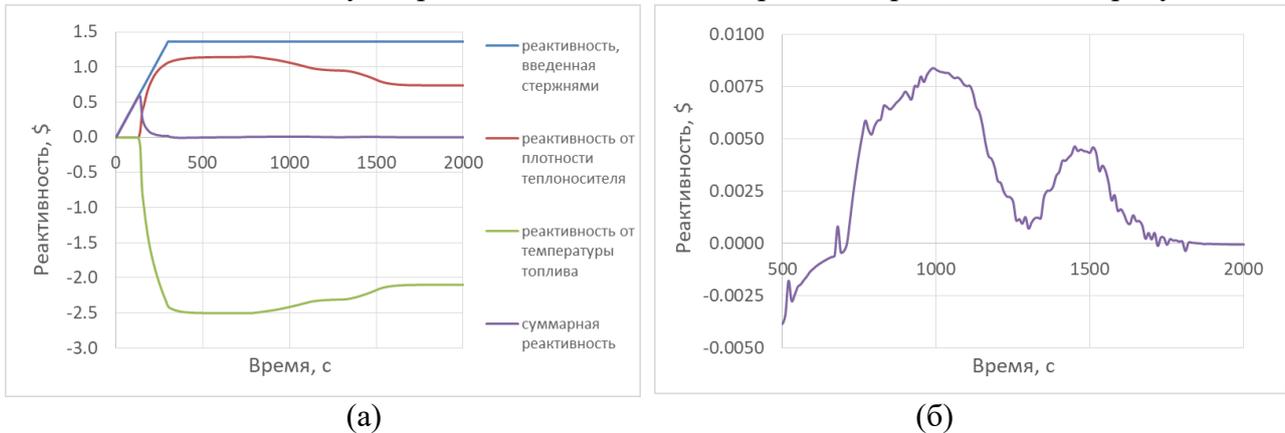


Рис. 5. Изменения учитываемых эффектов реактивностей: (а) – составляющие реактивности в зависимости от времени; (б) – суммарная реактивность

Приведенные выше зависимости показывают, что модель приходит к устойчивому стационарному состоянию, соответствующему номинальному режиму ВВЭР-СКД. Данное состояние использовано в качестве начального для исследования переходных процессов, связанных с изменением параметров питательной воды на входе в парогенератор.

4.2. Изменение состояния реактора при снижении температуры питательной воды.

Рассматривалось снижение температуры питательной воды на 40 °С с номинального значения, равного 320 °С до 280 °С (рисунок 6а), соответствующее изменение мощности реактора показано на рисунке 6б.

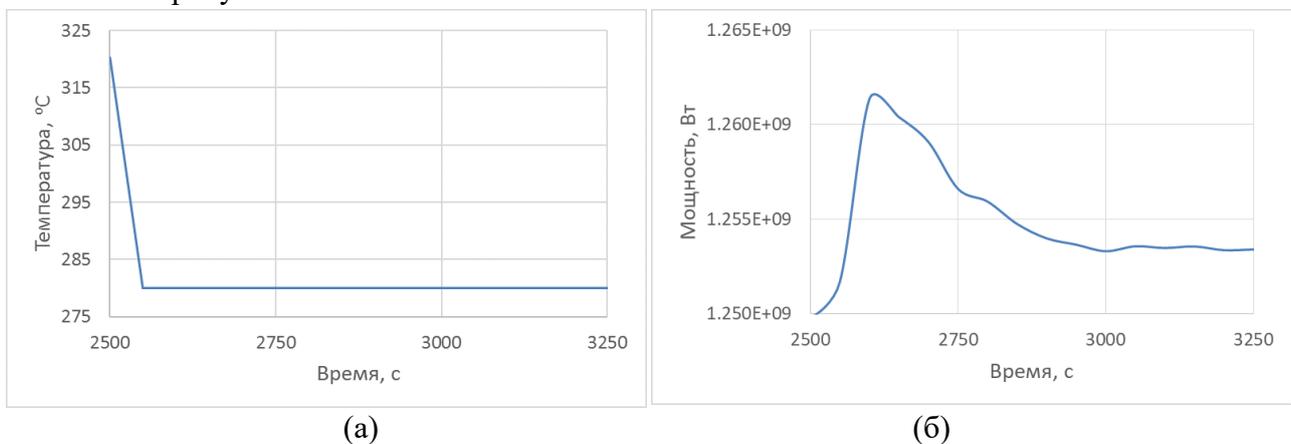


Рис. 6. Снижение температуры питательной воды до 280 °С (а); изменение мощности реактора в зависимости от времени (б)

Изменения температур топлива, оболочки ТВЭЛ и теплоносителя, а также реактивности представлены, соответственно, на рисунке 7а и рисунке 7б.

Из полученных зависимостей видно, что снижение температуры питательной воды на 40 °С при отсутствии вмешательства органов АСУ ТП приводит к незначительному (на 2 %) повышению мощности реактора (рисунок 6б), а также снижению на 5 °С температуры теплоносителя на входе в реактор до 400 °С (рисунок 7а) и на 40 °С температуры на выходе из реактора (до 480 °С). Температуры оболочек ТВЭЛ снижаются на 40 – 50 °С.

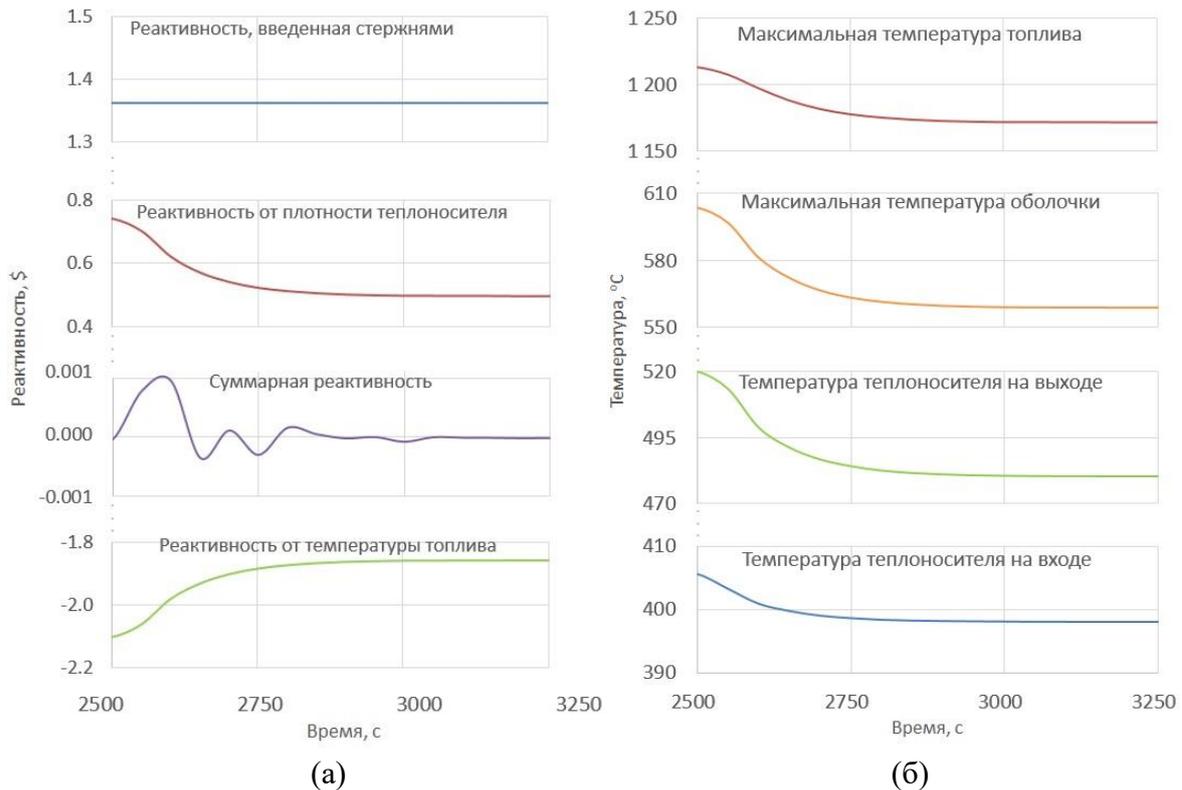


Рис. 7. Изменения максимальных температур топлива, оболочки и температуры теплоносителя на входе в реактор и выходе из него в зависимости от времени (а); составляющие реактивности в зависимости от времени (б) при снижении температуры питательной воды до 280 °С

4.3. Изменение состояния реактора при росте температуры питательной воды.

После случая снижения температуры питательной воды был рассмотрен случай ее повышения на 30 °С (с 320 °С до 350 °С), как показано на рисунке 8а. Соответствующее изменение мощности представлено на рисунке 8б.

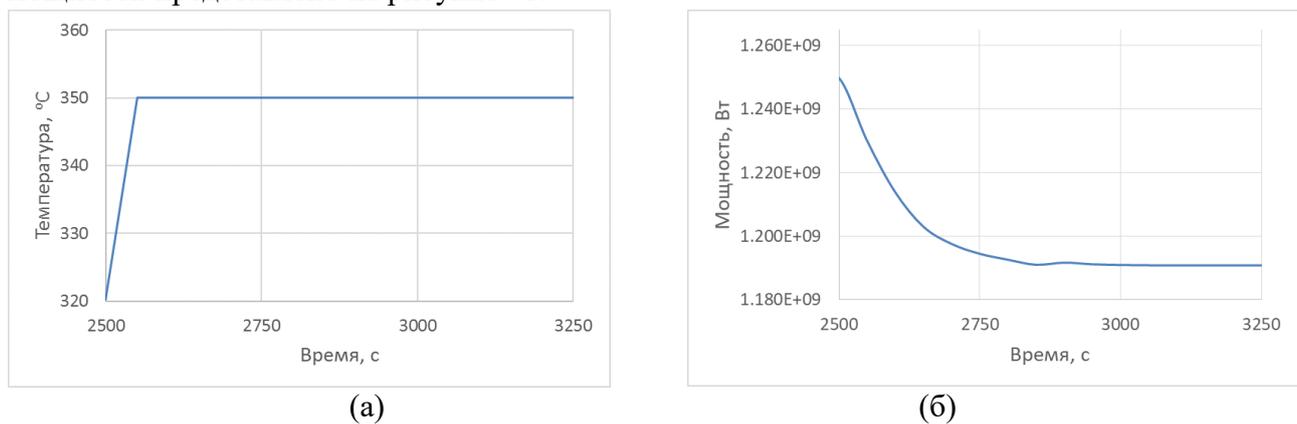


Рис. 8. Повышение температуры питательной воды до 350 °С (а); изменение мощности реактора в зависимости от времени (б)

Изменения максимальных температур топлива, оболочки ТВЭЛ и температур теплоносителя на входе в реактор и выходе из него представлены на рисунке 9а, а изменения составляющих реактивности представлены на рисунке 9б.

В данном случае из приведенных зависимостей видно, что при отсутствии срабатывания органов АСУ ТП мощность реактора снижается на 5 % (с 1250 МВт до 1190 МВт, рисунок 8б). При этом температура теплоносителя на входе в реактор повышается до 415 °С, а температура на выходе из реактора – до 545 °С, максимальные температуры оболочек ТВЭЛ повышаются примерно на 20 °С (рисунок 9а).

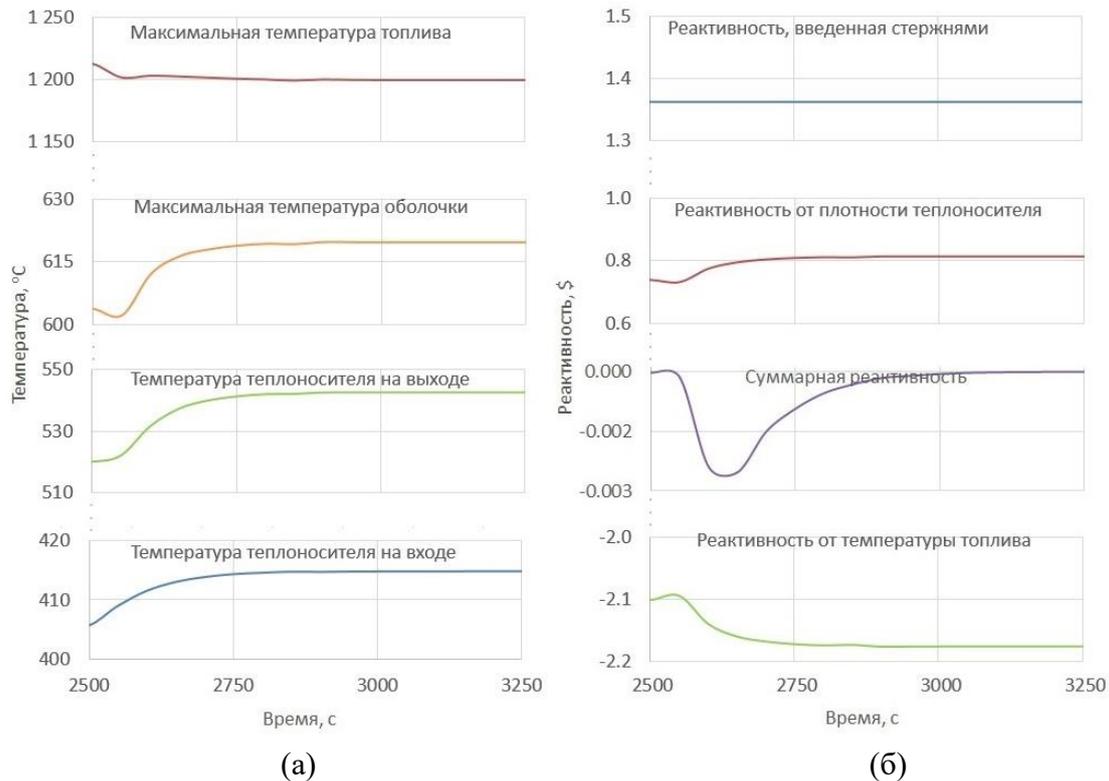


Рис. 9. Изменения максимальных температур топлива, оболочки и температуры теплоносителя на входе в реактор и выходе из него в зависимости от времени (а); составляющие реактивности в зависимости от времени (б) при повышении температуры питательной воды до 350 °С

4.4. Изменение состояния реактора при изменениях расхода питательной воды.

Исследовалось также влияние изменения расхода питательной воды на реакторную установку. В первом случае рассматривалось повышение расхода на 10 % (с 370 кг/с до 405 кг/с), что представлено на рисунке 10а.

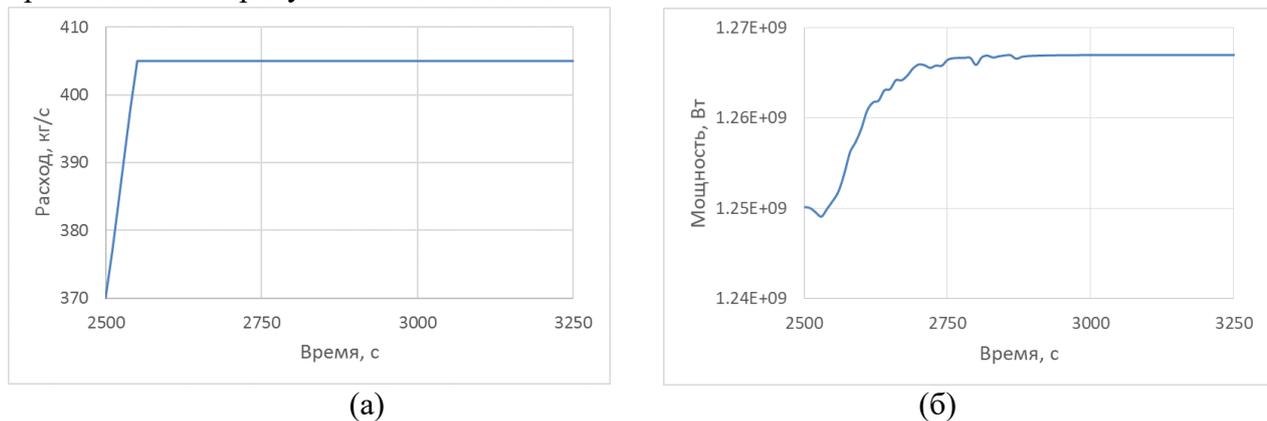


Рис. 10. Повышение массового расхода питательной воды до 405 кг/с (а); изменение мощности реактора в зависимости от времени (б)

Поведение мощности в данном переходном процессе показано на рисунке 10б. Изменения максимальных температур топлива, оболочки ТВЭЛ и температуры теплоносителя на входе/выходе реактора продемонстрированы на рисунке 11а, а изменение составляющих реактивности представлены на рисунке 11б.

Результаты, представленные на рисунках, показывают, что повышение расхода питательной воды на 10 % при отсутствии срабатывания органов АСУ ТП приводит к повышению мощности реактора на 1,6 % (с 1250 МВт практически до 1270 МВт, рисунок 10б). Температура теплоносителя на входе в реактор снижается на 4 °С (до 401 °С),

на выходе из реактора – на 19 °С (с 520 °С до 501 °С), максимальные температуры оболочек ТВЭЛов снижаются на 15 °С (рисунок 11а).

Следующим рассматриваемым случаем было снижение расхода питательной воды на 12,5 % до 324 кг/с, в соответствии с рисунком 12а. Изменения мощности реактора, соответствующие этому процессу, показаны на рисунке 12б. Зависимости максимальных температур топлива, оболочки ТВЭЛ и температуры теплоносителя на входе/выходе реактора показаны на рисунке 13а, а поведение составляющих реактивности в зависимости от времени для данного случая представлено на рисунке 13б.

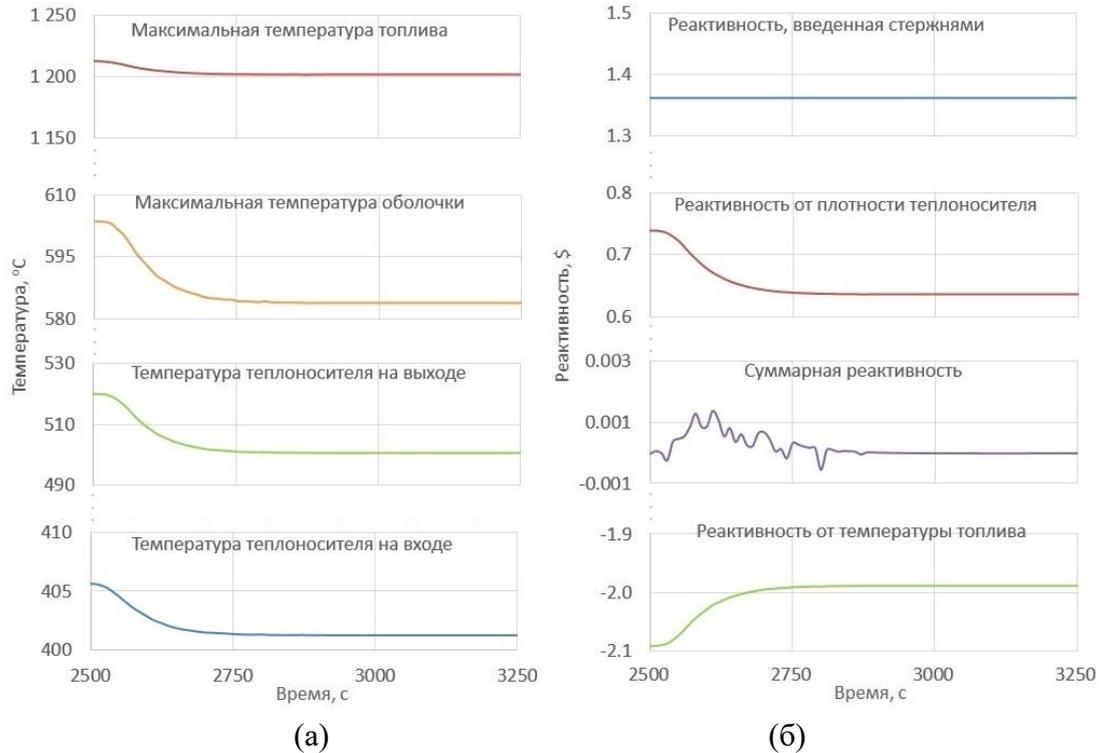


Рис. 11. Изменения максимальных температур топлива, оболочки и температуры теплоносителя на входе в реактор и выходе из него в зависимости от времени (а); составляющие реактивности в зависимости от времени (б) при повышении массового расхода питательной воды до 405 кг/с

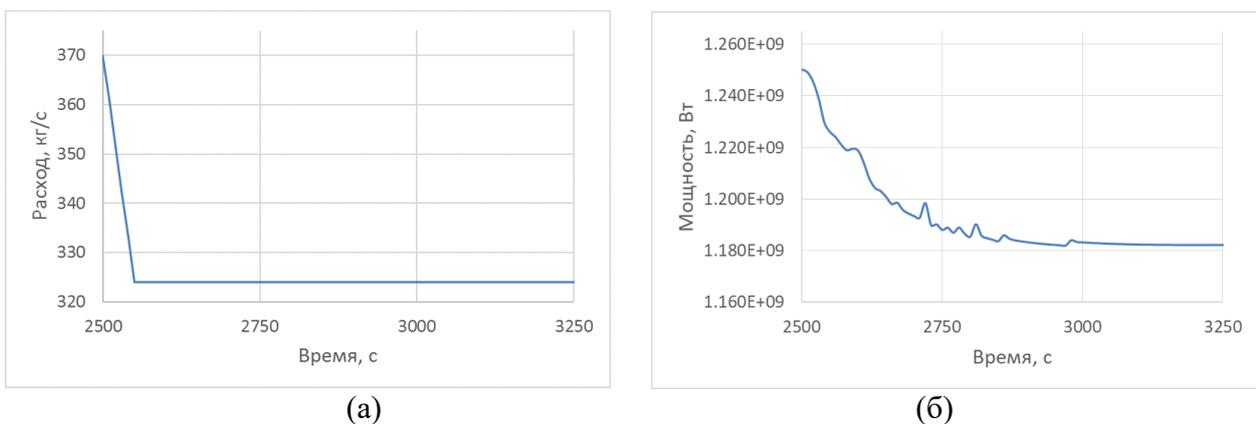


Рис. 12. Снижение массового расхода питательной воды до 324 кг/с (а); изменение мощности реактора в зависимости от времени (б)

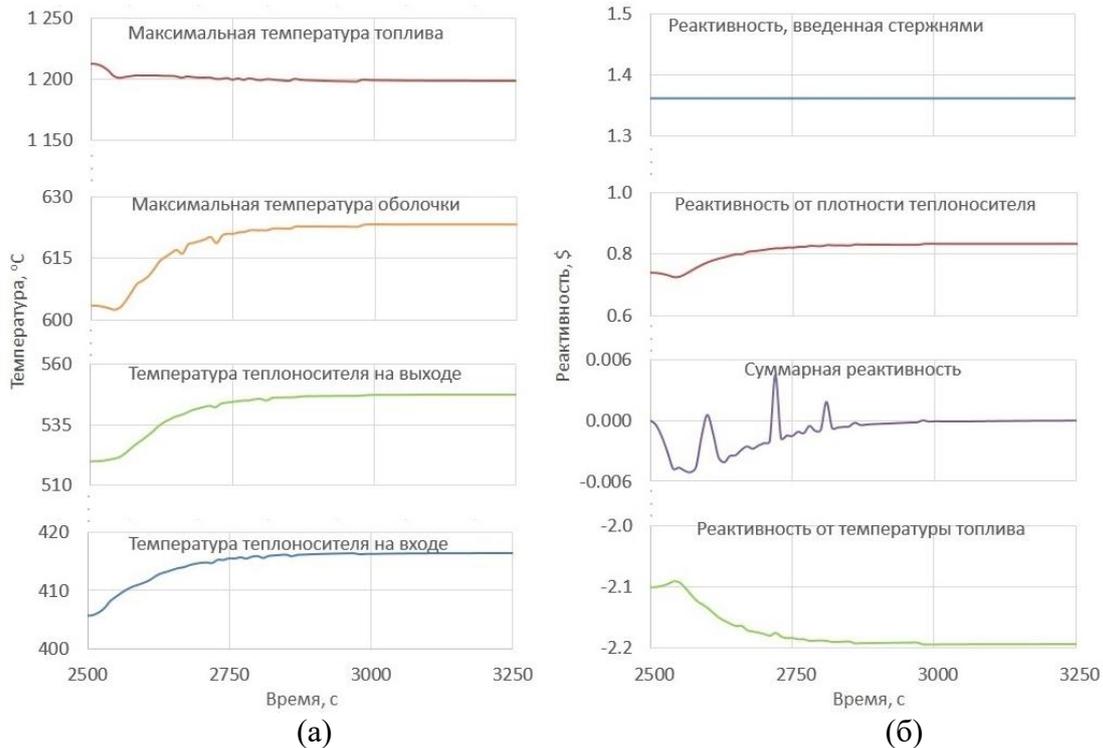


Рис. 13. Изменения максимальных температур топлива, оболочки и температуры теплоносителя на входе в реактор и выходе из него в зависимости от времени (а); составляющие реактивности в зависимости от времени (б) при снижении массового расхода питательной воды до 324 кг/с

Как видно из представленных рисунков, снижение расхода питательной воды на 12,5 % без работы органов АСУ ТП приводит к снижению мощности на 5,6 % (с 1250 МВт до 1180 МВт, рисунок 12б). Температура теплоносителя на входе в реактор при этом повышается на 10 °С (до 415 °С), а на выходе из реактора – на 25 °С (с 520 °С до 545 °С), значения максимальных температур оболочек ТВЭЛ повышаются на 20 °С (рисунок 13а).

5. Анализ результатов. Для расчета переходных режимов реакторной установки была разработана модель РУ ВВЭР-СКД с помощью программного кода RELAP5. С целью демонстрации работоспособности созданной модели, а также исследования динамики реактора, проведены расчеты переходных процессов. На первом этапе выполнен расчет условного пуска реактора по упрощенной схеме с линейным вводом реактивности регулирующих стержней реактора и последующим постепенным снижением температуры питательной воды до номинального уровня. Через 50 секунд после установления устойчивого номинального режима для разных расчетных случаев вносились следующие возмущения: снижение температуры питательной воды на 40 °С и ее повышение на 30 °С; повышение расхода питательной воды на 10 % и его снижение 12,5 %.

Проведенные расчеты показали, что в результате моделирования условного пуска система достигает устойчивого стационарного состояния, соответствующего номинальному режиму для РУ ВВЭР-СКД, что говорит об устойчивости системы реакторной установки к возмущениям со стороны питательной воды парогенератора.

Отклонение температуры питательной воды от номинального значения в диапазоне 280 – 350 °С не приводит к возникновению каких-либо нейтронно-теплогидравлических неустойчивостей. Мощность реактора без работы АСУ ТП устойчиво переходит на новый стационарный уровень (в диапазоне 95 – 102% от номинала). Температура теплоносителя на входе в реактор изменяется при этом в диапазоне 400 – 415 °С, температура на выходе из реактора – в диапазоне 480 – 545 °С, температуры оболочек ТВЭЛ – в диапазоне 550 – 620 °С.

Повышение расхода питательной воды на 10 % приводит к повышению мощности реактора на 1,6 % (с 1250 МВт до 1270 МВт). Температура теплоносителя на входе в реактор снижается на 4 °С (до 401 °С), на выходе из реактора – на 19 °С (с 520 °С до 501 °С). Максимальные температуры оболочек ТВЭЛ снижаются на 15 °С.

Снижение расхода питательной воды на 12,5 % без работы органов АСУ ТП приводит к снижению мощности на 5,6 % (с 1250 МВт до 1180 МВт). При этом температура теплоносителя на входе в реактор повышается на 10 °С (до 415 °С), на выходе из реактора – на 25 °С (с 520 °С до 545 °С). Максимальные температуры оболочек ТВЭЛов повышаются на 20 °С.

Заключение. В результате проведения расчетного исследования с применением RELAP5/MOD3.3 были получены зависимости, показывающие изменения основных параметров РУ ВВЭР-СКД во времени для случаев условного пуска реактора и переходных процессов, вызванных возмущениями параметров питательной воды на входе в парогенератор.

Расчет условного пуска реактора продемонстрировал выход системы на устойчивый номинальный режим работы, при котором мощность равна 1250 МВт, а температуры теплоносителя на входе в реактор и выходе из него равны 405 °С и 520 °С соответственно.

Исследование переходных процессов, при которых происходило снижение температуры питательной воды на входе в парогенератор на 40 °С, ее повышение на 30 °С, а также повышение расхода на 10 % и его снижение на 12,5 %, показало, что изменение температур оболочек не превышает при этом 20 %.

Таким образом, по итогу рассмотренных переходных процессов, все полученные значения температур находятся в допустимых пределах. В рассматриваемом диапазоне возмущений параметров питательной воды нейтронно-теплогидравлической неустойчивости не обнаружено.

Список источников

1. Locatelli G., Mancini M., Todeschini N. Generation IV nuclear reactors: Current status and future prospects. Energy Policy, 2013, pp. 1503-1520.
2. Duffey R.B., Pioro I.L. Supercritical water-cooled nuclear reactors: review and status. Nuclear materials and reactors from encyclopedia of life support systems (EOLSS), 2005, vol. 9, 14 p.
3. Sedov A.A., Pustovalov S.B., Shchepetina T.D., Simonov S.S. NPP with SCW nuclear power installation, in Proc. 10th Int. Symp. on SCWR (ISSCWR-10, Prague, Czech Republic, Mar. 15 – 19), 2021, p. 001.
4. Алексеев П.Н. Развитие технологии ВВЭР на базе энергетических реакторов с легководным теплоносителем сверхкритических параметров / П.Н. Алексеев, А.А. Ковалишин, А.А. Седов и др. // Вопросы атомной науки и техники. Серия: Физика ядерных реакторов, 2023. – №1. – С. 48-63.
5. Ambrosini W. On the analogies in the dynamic behaviour of heated channels with boiling and supercritical fluids. Nuclear Engineering and Design, 2007, v.237, pp. 1164-1174.
6. Украинцев В.Ф. Эффекты реактивности в энергетических реакторах / В.Ф. Украинцев. – Учебное пособие, Обнинск, ИАТЭ, 2000. – 60 с.
7. Heat transfer behaviour and thermohydraulics code testing for supercritical water cooled reactors (SCWRs). IAEA-TECDOC-1746, 2014. – 510 p.
8. Buongiorno J., MacDonald P. Supercritical water reactor (SCWR). Progress report for the FY-03 Generation-IV R&D activities for the development of the SCWR in the US, INEEL/Ext-03-03-01210, INEEL, USA, 2003, 38 p.
9. Ishiwatari Y. et al. Improvements of feedwater controller for the Super Fast reactor, Journal of nuclear science and technology, 2010, vol.47, no.12, pp. 1155-1164.
10. Liu L., et al. SCWR transient safety analysis code SCAC-CSR1000. Progress in nuclear energy, 2015, vol.83, pp. 318-325.
11. The RELAP5 Development Team. RELAP5/MOD3.3 Code manual Volume II: Appendix A Input Requirements. Information Systems Laboratories, Inc. Idaho Falls, Idaho, US Nuclear Regulatory Commission Washington DC, 2002, 20555, 306 p.
12. Lou M. Loss of coolant accident simulation for the Canadian supercritical water-cooled reactor using RELAP5/MOD4, Master thesis., 2016, 159 p.

13. NIST Reference fluid thermodynamic and transport properties database (REFPROP): Version 10. National Institute of Standards and Technology (NIST), U.S. Department of Commerce, available at: <https://www.nist.gov/srd/refprop> (accessed: 07/02/2023).

Суджян Артавазд Манукович. Научный сотрудник НИЦ КИ. AuthorID: 1133311, SPIN: 2947-4122, ORCID: 0000-0001-6943-9220, artavazd1994@gmail.com, 123098, г. Москва, пл. Академика Курчатова, 1.

UDC 536.24.083

DOI:10.25729/ESI.2023.32.4.009

Investigation of the dynamics of the VVER-SKD reactor plant with various changes in feed water parameters

Artavazd M. Sujyan

National Research Center “Kurchatov Institute”,
Russia, Moscow, artavazd1994@gmail.com

Abstract. This article presents the results of computational work on the study of the dynamics of the VVER-SKD reactor installation, performed using the capabilities of the RELAP5/MOD3.3 program. A characteristic feature of this reactor is a rather noticeable change in the density of the coolant in the core. In combination with neutron-physical feedback, it creates risks of thermal-hydraulic and neutron-thermal-hydraulic instability, especially during transients. In this regard, it is necessary to be able to predict the occurrence of instabilities, to assess the limits of the stability of the system. For these purposes, a design model of the reactor plant was developed. The choice of the RELAP5 code for this work is due to the extensive experience of its use in justifying the safety of existing water-cooled reactors. To obtain the possibility of calculations of VVER-SKD by the RELAP5 code, the properties of water were expanded and detailed in the field of supercritical pressures and temperatures. For the developed model of the VVER-SKD reactor plant, calculations of the conditional start-up of the reactor to the nominal operating mode were performed. The influence of deviations of the feed water temperature and flow rate on the dynamics of the reactor plant was also considered. Based on the results of calculations, conclusions are made about the stability of the reactor plant in the nominal mode.

Keywords: reactor plant, VVER-SKD, dynamics, stability, feedwater, feedbacks

References

1. Locatelli G., Mancini M., Todeschini N. Generation IV nuclear reactors: Current status and future prospects. *Energy Policy*, 2013, pp. 1503-1520.
2. Duffey R.B., Pioro I.L. Supercritical water-cooled nuclear reactors: review and status. *Nuclear materials and reactors from encyclopedia of life support systems (EOLSS)*, 2005, vol. 9, 14 p.
3. Sedov A.A., Pustovalov S.B., Shchepetina T.D., Simonov S.S. NPP with SCW nuclear power installation, in Proc. 10th Int. Symp. on SCWR (ISSCWR-10, Prague, Czech Republic, Mar. 15 – 19), 2021, p. 001.
4. Alekseev P.N., Kovalishin A.A., Sedov A.A., et al. Razvitiye tekhnologii VVER na baze energeticheskikh reaktorov s legkovodnym teplonositelem sverkhkriticheskikh parametrov [Development of VVER technology based on power reactors with light-water coolant of supercritical parameters]. *Voprosy atomnoy nauki i tekhniki. Seriya: Fizika yadernykh reaktorov [Problems of atomic science and engineering, Ser. Physics of Nuclear Reactors]*, 2023, iss. 1, pp. 48-63.
5. Ambrosini W. On the analogies in the dynamic behaviour of heated channels with boiling and supercritical fluids, *Nuclear Engineering and Design*, 2007, v.237, pp. 1164-1174.
6. Ukrainsev V.F. *Effekty reaktivnosti v energeticheskikh reaktorakh [Reactivity effects in power reactors]*. Obninsk: IATE, 2000, 60 p.
7. Heat transfer behaviour and thermohydraulics code testing for supercritical water cooled reactors (SCWRs). IAEA-TECDOC-1746, 2014. – 510 p.
8. Buongiorno J., MacDonald P. Supercritical water reactor (SCWR). Progress report for the FY-03 Generation-IV R&D activities for the development of the SCWR in the US, INEEL/Ext-03-03-01210, INEEL, USA, 2003, 38 p.

9. Ishiwatari Y. et al. Improvements of feedwater controller for the Super Fast reactor, Journal of nuclear science and technology, 2010, vol.47, no.12, pp. 1155-1164.
10. Liu L., et al. SCWR transient safety analysis code SCAC-CSR1000. Progress in nuclear energy, 2015, vol.83, pp. 318-325.
11. The RELAP5 Development Team. RELAP5/MOD3.3 Code manual Volume II: Appendix A Input Requirements. Information Systems Laboratories, Inc. Idaho Falls, Idaho, US Nuclear Regulatory Commission Washington DC, 2002, 20555, 306 p.
12. Lou M. Loss of coolant accident simulation for the Canadian supercritical water-cooled reactor using RELAP5/MOD4, Master thesis., 2016, 159 p.
13. NIST Reference fluid thermodynamic and transport properties database (REFPROP): Version 10. National Institute of Standards and Technology (NIST), U.S. Department of Commerce, available at: <https://www.nist.gov/srd/refprop> (accessed: 07/02/2023).

Sujyan Artavazd Manukovich. *Researcher at SRC KI. AuthorID: 1133311, SPIN: 2947-4122, ORCID: 0000-0001-6943-9220, artavazd1994@gmail.com, 123098, Moscow, square. Academician Kurchatov, 1.*

Статья поступила в редакцию 10.08.2023; одобрена после рецензирования 05.10.2023; принята к публикации 16.12.2023.

The article was submitted 08/10/2023; approved after reviewing 10/05/2023; accepted for publication 12/16/2023.

УДК 519.873+621.311

DOI: 10.25729/ESI.2023.32.4.010

Методика оценки живучести автономной микросети

Еделев Алексей Владимирович¹, Карамов Дмитрий Николаевич²,

Башарина Ольга Юрьевна^{3,4}

¹Институт систем энергетики им. Л.А. Мелентьева СО РАН, Россия, Иркутск,

flower@isem.irk.ru

²Иркутский национальный исследовательский технический университет, Россия, Иркутск

³Институт динамики систем и теории управления им. В.М. Матросова СО РАН, Россия,

Иркутск

⁴Уральский государственный экономический университет, Россия, Екатеринбург

Аннотация. Данная статья является второй из цикла, посвящённого исследованию живучести изолированных энергетических комплексов локального уровня или автономных микросетей с помощью ранее разработанной технологии цифрового двойника сложной технической системы. Живучестью называется свойство микросетей адаптироваться к крупным возмущениям и восстанавливать свое исходное состояние после их воздействия. Исследование живучести энергетических комплексов традиционно строится на многовариантных вычислительных экспериментах. Однако, цифровой двойник, связанный с реальной микросетью или испытательным стендом, позволяет в исследовании живучести микросетей комбинировать вычислительные эксперименты и натурные опыты. Двухстороннюю связь цифрового двойника с оборудованием микросети или стенда обеспечивает специализированная предметно-ориентированная среда, архитектура которой представлена в данной статье. Предлагаемая архитектура среды включает в себя систему мониторинга, которая, помимо сбора данных о состоянии вычислительных средств и коммуникационного оборудования, адаптирована к сбору данных с контрольно-измерительных приборов силового оборудования и автоматики микросети. В статье также представлена авторская методика оценки живучести автономной микросети с помощью ее цифрового двойника. Входными данными для оценки живучести, согласно этой методике, являются значения параметров цифрового двойника, информация из системы мониторинга, конфигурации микросети, показатели производительности, сводные показатели. На выходе модели строятся кривые живучести. Разработанная методика может использоваться при решении различных классов задач предметной области исследования живучести, например, в анализе уязвимости микросетей.

Ключевые слова: микросеть, живучесть, уязвимость, математическая модель, энергетический хаб, предметно-ориентированная среда, испытательный стенд

Цитирование: Еделев А.В. Методика оценки живучести автономной микросети / А.В. Еделев, Д.Н. Карамов, О.Ю. Башарина // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С.117-126. – DOI: 10.25729/ESI.2023.32.4.010.

Введение. Автономные системы энергоснабжения являются разновидностью энергетических комплексов локального уровня (микросетей), в которых совместное использование возобновляемых источников энергии с традиционными энергетическими установками и накопителями электрической энергии является экономически эффективным способом энергоснабжения потребителей. Особую значимость и большое распространение возобновляемые источники энергии получили при электрификации труднодоступных территорий, не имеющих связи с традиционными топливно-энергетическими комплексами [1].

Изолированность автономных микросетей повышает значимость исследования их функционирования в следующих экстремальных условиях:

- природные бедствия, такие как ураганы, наводнения, землетрясения, резкие похолодания и т.д.;
- отказ элементов и подсистем вследствие износа, ошибок персонала и т.д.;
- преднамеренные (умышленные) нарушения, такие как кибератаки и т.д.

Способность противостоять вышеперечисленным крупным возмущениям, не допуская каскадного развития аварий с массовым нарушением режима энергоснабжения потребителей, и восстанавливать исходное состояние или близкое к нему характеризует одно из свойств энергетических комплексов – живучесть [2]. Традиционно исследование живучести строится на проведении масштабных вычислительных экспериментов, в ходе которых разыгрывается множество сценариев крупных возмущений с применением математических моделей [3].

В работе [4] начато обсуждение возможности применения цифрового двойника, как средства для исследования свойств моделируемого объекта, для исследования живучести микросетей. Ключевым компонентом цифрового двойника микросети является комплекс оптимизационных, имитационных, нейросетевых моделей, с высокой точностью описывающих все аспекты поведения этого энергетического комплекса [5]. Однако, отличие цифрового двойника от модели объекта исследования заключается в наличии обратной связи с объектом, что позволяет в процессе исследования гибко подстраивать управление моделируемым объектом, исходя из разницы между прогнозируемыми (полученными в результате расчета на математической модели) и реальными значениями технологических параметров оборудования. Следует отметить, что оборудование современных испытательных стендов (рис. 1) позволяет разворачивать полноценные реалистичные схемы микросетей и имитировать воздействие на них крупных возмущений. Таким образом, цифровой двойник, соединенный с испытательным стендом, даёт возможность сочетать в исследовании живучести микросети вычислительные эксперименты и натурные опыты.

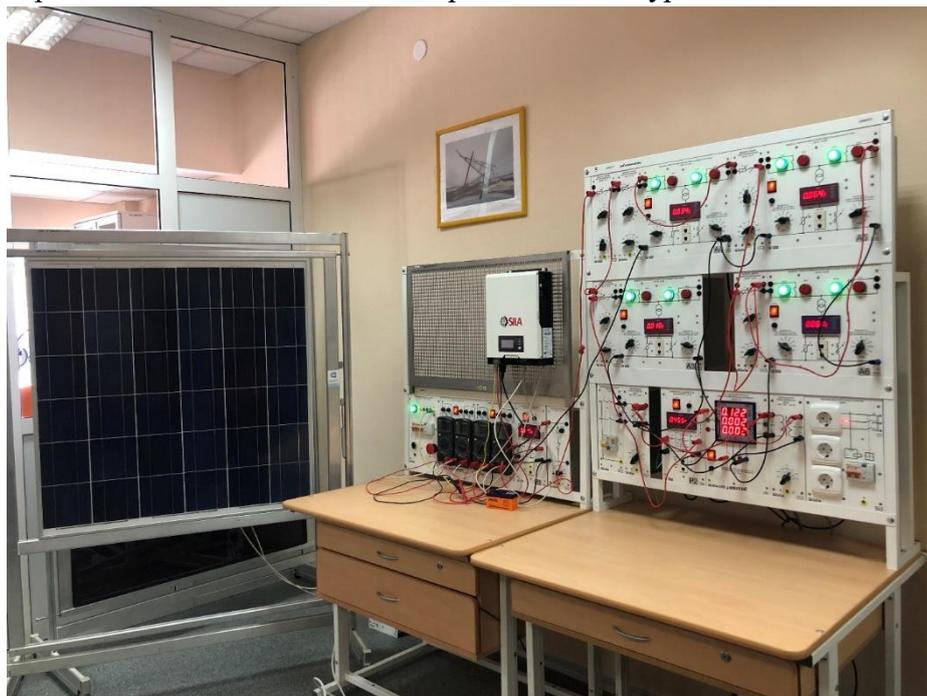


Рис. 1. Испытательный стенд, имитирующий автономную микросеть

Комплекс моделей для цифрового двойника микросети был представлен в предыдущей работе [4]. Цель данной статьи – представить архитектуру предметно-ориентированной среды, в которой функционирует цифровой двойник микросети, и разработанную методику оценки живучести автономной микросети с помощью цифрового двойника.

1. Предметно-ориентированная среда для исследования живучести автономных микросетей. Под предметно-ориентированной средой понимается совокупность программно-аппаратных средств, позволяющих её конечным пользователям решать один или несколько определённых классов задач некоторой предметной области [6, 7].

Методическую основу предметно-ориентированной среды для исследования живучести энергетических комплексов составляют:

- единые принципы взаимодействия с моделями энергетических комплексов разных уровней территориально-отраслевой иерархии [8];
- механизм объединения различных моделей отдельных систем энергетики для исследования их совместной работы в экстремальных условиях [9];
- отображение крупных возмущений на модель энергетических комплексов путём моделирования отказов и восстановлений отдельных элементов системы;
- разделение крупных возмущений по типу воздействия на структурные и функциональные;
- возможность моделирования каскадного развития аварий, в том числе их распространения по взаимосвязанным системам энергетики.

Состав предметно-ориентированной среды для исследования живучести энергетических комплексов подробно описан в работе [3]. Главными компонентами этой среды являются:

- распределённые пакеты прикладных программ (РППП), отвечающие за решение определённого класса задач [10];
- средства создания автоматизированных рабочих мест, которые позволяют исследователю работать с исходными данными, взаимодействовать с РППП, анализировать результаты вычислительных экспериментов [11].

РППП разрабатываются и выполняются под управлением инструментария Orlando Tools [12]. В отличие от известных инструментариев, средства Orlando Tools базируются на применении инженерии знаний, непрерывной интеграции системного и прикладного программного обеспечения, конфигурирования ресурсов и мультиагентного управления. Ускорение вычислений, обеспечение их надёжности и масштабируемости, а также повышение эффективности использования ресурсов при решении больших задач достигаются путём смягчения неопределённостей, возникающих при распределении вычислительной нагрузки, за счёт извлечения знаний о времени выполнения приложений на ресурсах, классификации заданий, использования знаний о соответствии ресурсов классам заданий, а также сведений, получаемых в процессе разработки, доставки, развёртывания и тестирования программного обеспечения.

Как говорилось во введении, предметно-ориентированная среда для исследования живучести автономной микросети должна быть ориентирована на проведение не только вычислительных экспериментов, но и натурных опытов на энергетическом оборудовании. Поэтому в архитектуру среды (рис. 2) включена система мониторинга, которая, помимо сбора данных о состоянии вычислительных средств и коммуникационного оборудования, адаптирована к сбору данных с контрольно-измерительных приборов силового оборудования и автоматики микросети [13]. Перечисленное оборудование может входить в состав испытательного стенда, имитирующего реальную автономную микросеть (рис. 1). Также система мониторинга обеспечивает представление ретроспективной природно-климатической информации, извлечённой из циклической базы данных, в виде типичного метеорологического года или многолетних временных рядов.

Цифровой двойник автономной микросети реализуется в виде набора веб-сервисов [14, 15], организация и взаимодействие которых основывается на микро-сервисном подходе [16]. Сервисы, перенаправляя вызовы системе мониторинга и необходимым РППП (рис. 2), обеспечивают возможность проведения массовых расчетов:

- по агрегированию, структурированию и анализу ретроспективных и текущих предметных данных большого объема, включая выявление и прогнозирование трендов,

- а также регулярных (сезонных и суточных) колебаний и случайных отклонений в динамике изменений природно-климатических показателей [17, 18];
- по моделированию процесса управления работой оборудования [19];
- для анализа эффективности и оптимизации энергоснабжения [20];
- для исследования живучести [4], включая анализ уязвимости [21, 22].

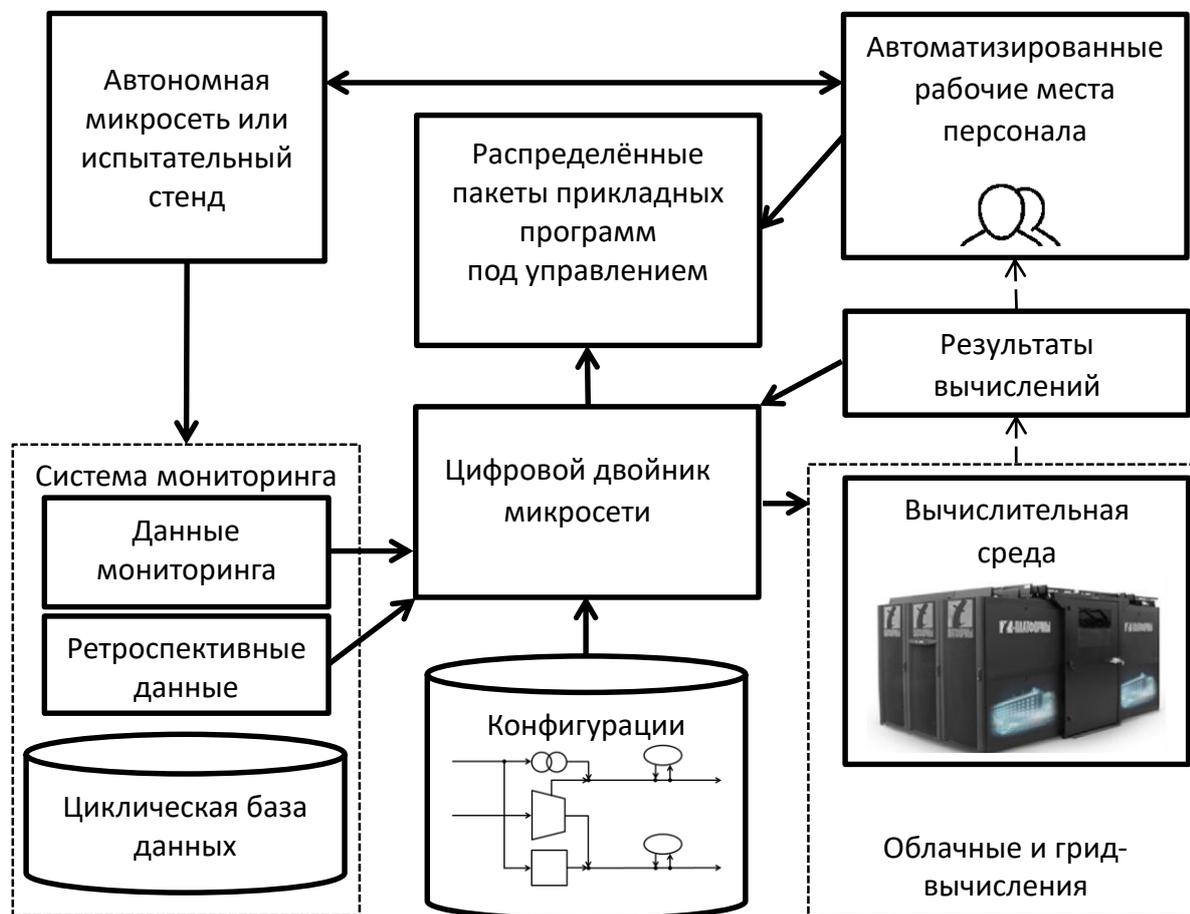


Рис. 2. Архитектура предметно-ориентированной среды для исследования живучести автономной микросети

2. Оценка живучести автономной микросети. Как было отмечено во введении, с точки зрения надёжного бесперебойного энергоснабжения потребителей изолированных районов наиболее актуальной задачей является оценка эффективности выполнения автономной микросетью своих функций в экстремальных условиях функционирования. Сценарная оценка живучести является одним из таких инструментов, она отслеживает и исследует изменения между состояниями энергетического комплекса до, во время и после воздействия крупного возмущения [23].

Как показано в предыдущем разделе, в предметно-ориентированной среде для исследования живучести источники (механизмы генерации) временных рядов данных, представляющих состояния энергетического комплекса в последовательные моменты времени сценария крупного возмущения, могут варьироваться в широких пределах: исторические данные, натурные опыты, агентное и имитационное моделирование системной динамики [24], искусственные нейронные сети [25], сетевые потоковые методы [26]. К последним относится описанная в работе [4] модель микросети на основе концепции энергетического хаба.

Цифровой двойник (рис. 2) каким-либо из перечисленных способов генерирует n временных рядов данных, представляющих значения параметров энергетического комплекса x_1, x_2, \dots, x_n . Выбор способа генерации рядов данных и состава вектора состояния $\mathbf{x} = (x_1, x_2, \dots, x_n)$ определяется целями сторон, заинтересованных в исследовании живучести, например производителей и потребителей энергоресурсов, оперативного персонала и т.д. [23]. В состав вектора \mathbf{x} входят технологические, социальные и экономические параметры объектов энергетики, а также природно-климатические параметры, которые представляют условия внешней среды.

Если через X обозначить последовательность состояний \mathbf{x} в последовательные моменты времени $t = \{1, \dots, T\}$, то можно ввести понятие показателя производительности $P(t): X \rightarrow \mathbb{R}$. Эта функция отображает в скалярную величину состояние $\mathbf{x} \in X$, например, возвращая значение одного из параметров энергетического комплекса x_1, x_2, \dots, x_n в момент времени t . Для сравнения конфигураций энергетического комплекса по живучести применяется сводный показатель $J: P \rightarrow \mathbb{R}$, который сводит временной ряд P к скалярной величине.

График показателя производительности P называется кривой живучести. Набор кривых живучести наглядно иллюстрирует поведение энергетического комплекса при крупном возмущении и позволяет количественно оценить его живучесть.

Сценарная оценка живучести автономной микросети, представленная на рис. 3, разработана на базе общей схемы поддержки исследования живучести энергетических комплексов, предложенной в работе [3]. Входными данными для оценки живучести являются значения параметров цифрового двойника, информация из системы мониторинга, конфигурации микросети, показатели производительности, сводные показатели, а на выходе строятся кривые живучести.



Рис. 3. Методика оценки живучести автономной микросети

На первом шаге методики оценки живучести автономной микросети (рис. 3) при формировании множества сценариев крупных возмущений при большом объёме данных могут использоваться возможности цифрового двойника по анализу временных рядов данных для выявления наиболее тяжёлых природно-климатических и социально-экономических условий функционирования микросети [17, 18]. Найденные экстремальные условия закладываются в сценарии крупных возмущений, которые моделируют одиночные или групповые отказы отобранных объектов энергетики [21, 22].

На втором шаге для каждого сценария из сформированного множества с помощью цифрового двойника проводится генерация последовательностей состояний микросети.

На следующем шаге методики сгенерированные последовательности состояний автономной микросети используются для расчета и нормализации заданных показателей производительности и сводных показателей. Нормализация показателей обоих видов может производиться статическим, эндогенным или экзогенным способами, недостатки и преимущества которых обсуждаются в работе [23].

На последнем шаге оценки живучести автономной микросети по рассчитанным показателям производительности строятся кривые живучести.

Результаты сценарной оценки живучести: кривые живучести и значения сводных показателей используются в различных классах задач данной предметной области, например, в анализе уязвимости [21, 22].

Заключение. Выше была рассмотрена архитектура предметно-ориентированной среды для исследования живучести автономных микросетей. В целом, под подобной средой понимается совокупность программно-аппаратных средств, позволяющих её конечным пользователям решать один или несколько определённых классов задач некоторой предметной области. В данной работе решаемым классом задач была выбрана оценка живучести энергетических комплексов локального уровня, или микросетей.

Особенностью представленной в статье методики оценки живучести является ее направленность на использование цифрового двойника, который имеет двухстороннюю связь с испытательным стендом. С одной стороны, это позволяет при исследовании поведения энергетических комплексов применять не только вычислительные эксперименты, но и натурные опыты. Также цифровой двойник по сравнению с отдельной моделью объекта исследования имеет обратную связь с изучаемым объектом, что позволяет в динамике влиять на поведение микросети, исходя из разницы между прогнозируемыми и реальными значениями технологических параметров оборудования. С другой стороны, эта особенность и моделирование микросети на основе концепции энергетического хаба ограничивают применимость представленной методики энергетическими комплексами локального уровня, которые в силу экстремальных условий потеряли соединение с внешними системами энергетики, либо осуществляют энергоснабжение отдаленных или труднодоступных территорий.

Дальнейшее развитие данной работы видится в практическом применении описанной методики оценки живучести автономных микросетей.

Благодарности. Исследование выполнено при поддержке Министерства науки и высшего образования Российской Федерации, проект № FWEW-2021-0005 «Технологии разработки и анализа предметно-ориентированных интеллектуальных систем группового управления в недетерминированных распределенных средах» с использованием ресурсов центров публичного доступа «Иркутский суперкомпьютерный центр СО РАН» (<http://hpc.icc.ru>) и «Сибирский суперкомпьютерный центр СО РАН» (<http://www.sccc.icmmg.nsc.ru>).

Список источников

1. Карамов Д.Н. Интеграция процесса категоризации электрохимических накопителей энергии в задачу оптимизации состава оборудования автономных энергетических комплексов, использующих возобновляемые источники энергии / Д.Н. Карамов // Известия Томского политехнического университета. Инжиниринг георесурсов, 2019. – Т. 330. – № 5. – С. 113-130.
2. Надежность систем энергетики и их оборудования. В 4-х томах. Т. 1. Справочник по общим моделям анализа и синтеза надежности систем энергетики / ред. Ю.Н. Руденко. – М.: Энергоатомиздат. –1994. – 480 с.
3. Vyckov I.V., Gorsky S.A., Edelev A.V. et.al. Support for managing the resilience of energy systems based on a combinatorial approach. Journal of computer and systems sciences international, 2021, v. 60, no. 6, pp. 981-994, DOI: 10.1134/S1064230721060071.

4. Еделев А.В. Моделирование автономной микросети / А.В. Еделев, Д.Н. Карамов, О.Ю. Башарина // Информационные и математические технологии в науке и управлении, 2023. – № 3(31). – С.74-85. – DOI:10.25729/ESI.2023.31.3.007.
5. Тягунов М.Г. Использование цифрового двойника гибридных энергетических комплексов для оптимизации их параметров и режимов / М.Г. Тягунов, Р.П. Шевердиев // Вестник МЭИ, 2023. – № 3. – С. 109-118. – DOI: 10.24160/1993-6982-2023-3-109-118.
6. Gallopoulos E., Houstis E., Rice J.R. Computer as Thinker/Doer: Problem-solving environments for computational science. IEEE Computational science and engineering 1.2, 1994, pp. 11-23, DOI: 10.1109/99. 326669.
7. Houstis E., Gallopoulos E., Bramley R., Rice J. Problem-solving environments for computational science. IEEE Computational science and engineering, 1997, v. 34, no. 3, pp. 18-21.
8. Еделев А.В. Выбор подхода к моделированию взаимосвязанных критических инфраструктур / А.В. Еделев, А.Г. Феоктистов // Современные проблемы и перспективные направления инновационного развития науки: Тр. Международной научно-практической конф., 2021. – Калуга: Аэтерна. – С. 25–30. – URL: <https://aeterna-ufa.ru/sbornik/NK-345.pdf#page=25> (дата обращения 28.05.2023).
9. Еделев А.В. Подход к моделированию функционирования взаимосвязанных систем энергетики в условиях возмущений и его программная поддержка / А.В. Еделев, Н.М. Береснева // Программные продукты и системы, 2021. – Т. 34. – № 3. – С. 409–419. – DOI: 10.15827/0236-235X.135.409-419.
10. Феоктистов А.Г. Инструментальные средства разработки распределенных пакетов прикладных программ на основе модульного программирования / А.Г. Феоктистов., И.А. Сидоров, С.А. Горский // Марчужковские научные чтения, 2017. – С. 950-956.
11. Фереферов Е.С. Технология разработки приложений баз данных на основе декларативных спецификаций / Е.С. Фереферов, И.В. Бычков, А.Е. Хмельнов // Вычислительные технологии, 2014. – Т. 19. – № 5. – С. 85-100.
12. Tchernykh A., Bychkov I., Feoktistov A. et.al. Mitigating uncertainty in developing and applying scientific applications in an integrated computing environment. Programming and computer software, 2020, 46, 8, pp. 483-502.
13. Sidorov I., Kostromin R., Feoktistov A. System for monitoring parameters of functioning infrastructure objects and their external environment. Proceedings of the 2nd International workshop on information, computation, and control systems for distributed environments, CEUR-WS Proceedings, 2020, v. 2638, pp. 252-264, DOI: 10.47350/ICCS-DE.2020.23.
14. Kostromin R., Feoktistov A., Voskoboinikov M. Service-oriented tools for automating digital twin development. Proceedings of the 4th Scientific-practical workshop on information technologies: algorithms, models, systems (ITAMS 2021), CEUR-WS Proceedings, 2021, v. 2984, pp. 95-100, DOI: 10.47350/ITAMS.2021.12.
15. Kostromin R., Feoktistov A. Agent-based DevOps of software and hardware resources for digital twins of infrastructural objects. Proceedings of the 4th International conference on future networks and distributed systems (ICFNDS 2020), ACM, 2020, pp. 1-6, DOI: 10.1145/3440749.
16. Kostromin R., Basharina O., Feoktistov A., Sidorov I. Microservice-based approach to simulating environmentally-friendly equipment of infrastructure objects taking into account meteorological data. Atmosphere, 2021, v. 12, no. 9, 1217, pp. 1-24, DOI: 10.3390/atmos12091217.
17. Башарина О.Ю. Методика выявления и прогнозирования составляющих временных рядов природно-климатических показателей / О.Ю. Башарина, А.Г. Феоктистов // Прорывные научные исследования как двигатель науки: Тр. Международной научно-практической конф, 2021. – Новосибирск: Аэтерна. – С. 20–24. – URL: <https://aeterna-ufa.ru/sbornik/NK-347.pdf#page=20> (дата обращения 28.05.2023).
18. Еделев А.В. Оценка трендов, регулярных сезонных и случайных колебаний природно-климатических характеристик с помощью среды R / А.В. Еделев // Способы, модели и алгоритмы модернизации науки в современных условиях: Тр. Международной научно-практической конф, 2020. – Уфа: Аэтерна. – С. 11-14. – URL: <https://aeterna-ufa.ru/sbornik/NK-318.pdf#page=11> (дата обращения 28.05.2023).
19. Феоктистов А.Г. Цифровые двойники процессов работы природосберегающего оборудования инфраструктурного объекта / А.Г. Феоктистов, Р.О. Костромин, И.А. Сидоров и др. // Современные наукоемкие технологии, 2021. – № 1. – С. 57-62. – DOI: 10.17513/snt.38471.
20. Карамов Д.Н. Моделирование энергоснабжения объектов охраняемой Байкальской природной территории на основе возобновляемых источников энергии / Д.Н. Карамов, А.В. Еделев, А.Г. Феоктистов // Международный технико-экономический журнал, 2020. – № 5. – С. 7-24. – DOI: 10.34286/1995-4646-2020-74-5-7-24.
21. Gorsky S., Edelev A., Feoktistov A. Data processing in problem-solving of energy system vulnerability based on in-memory data grid. Mathematics and its applications in new computer systems (MANCS 2021). Lecture notes in networks and systems, Springer, Cham, 2022, v. 424, pp. 271-279.

22. Еделев А.В. Анализ уязвимости энергетической инфраструктуры и его реализация / А.В. Еделев, Н.М. Береснева, С.А. Горский // Современные наукоемкие технологии, 2022. – № 1. – С. 47-52. – DOI: 10.17513/snt.39008.
23. Poulin C.R., Kane M.B. Infrastructure resilience curves: Performance measures and summary metrics. Reliability engineering & system safety, 2021, v. 216, p. 107926, DOI: 10.1016/j.res.2021.107926.
24. Liu X., Fang Y.P., Zio E. A hierarchical resilience enhancement framework for interdependent critical infrastructures. Reliability engineering & system safety, 2021, v. 215, p. 107868.
25. Liu X., Ferrario E., Zio E. Identifying resilient-important elements in interdependent critical infrastructures by sensitivity analysis. Reliability engineering & system safety, 2019, v. 189, p. 423-434.
26. Johansson J., Hassel H. Modelling, simulation and vulnerability analysis of interdependent technical infrastructures. Risk and interdependencies in critical infrastructures: a guideline for analysis, London, Springer-Verlag, 2012, pp. 49-66.

Еделев Алексей Владимирович. К.т.н., старший научный сотрудник Института систем энергетики им. Л.А. Мелентьева. Область научных интересов – живучесть энергетических комплексов, математическое моделирование, распределенные вычисления. AuthorID: 710350, SPIN: 8447-9522, ORCID: 0000-0003-2219-9754. flower@isem.irk.ru, Россия, г. Иркутск, Лермонтова, 130.

Карамов Дмитрий Николаевич. К.т.н., доцент, Иркутский национальный исследовательский технический университет, Байкальский институт БРИКС. Область научных интересов – энергетические системы и комплексы, возобновляемые источники энергии, математическое моделирование. AuthorID: 905773, SPIN: 9977-1303, ORCID: 0000-0001-5360-4826, dmitriy.karamov@mail.ru, Россия, г. Иркутск, Лермонтова, 83.

Башарина Ольга Юрьевна. К.т.н., доцент, научный сотрудник Института динамики систем и теории управления им. В.М. Матросова СО РАН, доцент Уральского государственного экономического университета. Область научных интересов – системный анализ и моделирование сложных систем. AuthorID: 702499, SPIN: 2612-2891, ORCID: 0000-0002-7151-782X, basharinaolga@mail.ru, Россия, г. Иркутск, Лермонтова, 134.

UDC 519.873+621.311

DOI: 10.25729/ESI.2023.32.4.010

Methodology for assessing the resilience of an autonomous microgrid

Alexey V. Edelev¹, Dmitriy N. Karamov², Olga Yu. Basharina^{3, 4}

¹Melentiev Energy Systems Institute SB RAS, Russia, Irkutsk, flower@isem.irk.ru

²National Research Irkutsk State Technical University, Russia, Irkutsk

³Matrosov Institute for System Dynamics and Control Theory of SB RAS, Russia, Irkutsk

⁴Ural State University of Economics, Russia, Yekaterinburg

Abstract. This article is the second in a series devoted to the study of the resilience of isolated local-level energy complexes or autonomous microgrids using previously developed digital twin technology of a complex technical system. Vitality is the ability of microgrids to adapt to large disturbances and restore their original state after their impact. The study of the resilience of energy complexes is traditionally based on multivariate computational experiments. However, a digital twin associated with a real microgrid or test bench makes it possible to combine computational experiments and field experiments in the study of microgrid resilience. Two-way communication between a digital twin and the equipment of a microgrid or test bench is provided by a specialized subject-oriented environment, the architecture of which is presented in this article. The proposed environment architecture includes a monitoring system, which, in addition to collecting data on the state of computing facilities and communication equipment, is adapted to collecting data from instrumentation of power equipment and microgrid automation. The article also develops a methodology for assessing the resilience of an autonomous microgrid using its digital twin. The input data for the resilience assessment methodology are the values of the digital twin parameters, information from the monitoring system, microgrid configurations, performance indicators, summary indicators, and the output is resilience curves. The developed methodology will be further used in solving various classes of problems in the subject area of resilience research, for example, in analyzing the vulnerability of microgrids.

Keywords: microgrid, resilience, vulnerability, mathematical model, energy hub, subject-oriented environment, testbed

Acknowledgements: The study was supported by the Ministry of Science and Higher Education of the Russian Federation, project no. № FWEW-2021-0005 “Technologies for the development and analysis of subject-oriented intelligent group control systems in non-deterministic distributed environments”.

References

1. Karamov D.N. Integratsiya protsessa kategorizatsii elektrokhimicheskikh istochnikov energii v unikal'nom sostave oborudovaniya avtonomnykh energeticheskikh kompleksov, ispol'zuyushchikh istochniki energii [Integration of the storage battery categorization process into the task of optimizing the equipment of stand-alone energy systems with renewable energy sources]. *Izvestiya Tomskogo politekhnicheskogo universiteta. Inzhiniring georesursov* [Bulletin of the Tomsk Polytechnic University. Geo assets engineering], 2019, v. 330, no. 5, pp. 113-130.
2. Nadezhnost' sistem energetiki i ikh oborudovaniya [Reliability of energy systems and their equipment], vol. 1. Ed Rudenko Yu.N., Moscow, Energoatomizdat, 1994, 480 p.
3. Bychkov I.V., Gorsky S.A., Edelev A.V. et.al. Support for managing the resilience of energy systems based on a combinatorial approach. *Journal of computer and systems sciences international*, 2021, v. 60, no. 6. pp. 981-994, DOI: 10.1134/S1064230721060071.
4. Edelev A.V., Karamov D.N., Basharina O.Yu. Modelirovaniye avtonomnoy mikroseti [Modelling autonomous microgrids]. *Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii* [Information and mathematical technologies in science and management], 2023, no (31), pp.74-85, DOI:10.25729/ESI.2023.31.3.007.
5. Tyagunov M.G., Sheverdiyev R.P. Ispol'zovaniye dvoynika gibridnykh energeticheskikh kompleksov dlya optimizatsii ikh parametrov i rezhimov [Using a digital twin of hybrid energy complexes to optimize their parameters and modes]. *Vestnik MEI* [Bulletin of MPEI], no 3, pp. 109-118, DOI: 10.24160/1993-6982-2023-3-109-118.
6. Gallopoulos E., Houstis E., Rice J.R. Computer as Thinker/Doer: Problem-solving environments for computational science. *IEEE Computational science and engineering* 1.2, 1994, pp. 11-23, DOI: 10.1109/99. 326669.
7. Houstis E., Gallopoulos E., Bramley R., Rice J. Problem-solving environments for computational science. *IEEE Computational science and engineering*, 1997, v. 34, no. 3, pp. 18-21.
8. Edelev A.V., Feoktistov A.G. Vybor priverzhennosti modelirovaniyu ob'yedinennykh predprinimatel'skikh infrastruktur [Choosing an approach to modeling interconnected critical infrastructures]. *Trudy mezhdunarodnoy nauchno-prakticheskoy konferentsii “Sovremennyye problemy i perspektivnyye napravleniya innovatsionnogo razvitiya nauki”* [Proc. of International scientific and practical conference “Modern problems and promising directions of innovative development of science”], Kaluga, Aeterna, 2021, pp. 25-30, available at: <https://aeterna-ufa.ru/sbornik/NK-345.pdf#page=25> (accessed: 05/28/2023).
9. Edelev A.V., N.M. Beresneva Podkhod k modelirovaniyu integratsii sistem energetiki yego v usloviyakh vozmushcheniya i programmnoy podderzhki [An approach to modeling the functioning of interdependent energy systems under disturbances and its software support]. *Programmnyye produkty i sistemy* [Software & Systems], 2021, v. 34, no. 3, pp. 409-419, DOI: 10.15827/0236-235X.135.409-419.
10. Feoktistov A.G., Sidorov I.A., Gorsky S.A. Instrumental'nyye sredstva razrabotki rasprostranyayemykh paketov prikladnykh programm na osnove modul'nogo programmirovaniya [Development tools for distributed application packages based on modular programming]. *Marchukovskiye nauchnyye chteniya* [Marchuk scientific readings], 2017, pp. 950-956.
11. Fereferov E.S., Bychkov I.V., Hmel'nov A.E. Tekhnologiya razrabotki prilozheniy baz dannykh na osnove deklarativnykh spetsifikatsiy [Technology for database applications based on declarative specifications]. *Vychislitel'nyye tekhnologii* [Computational technologies], 2014, v. 19, no. 5, pp. 85-100.
12. Tchernykh A., Bychkov I., Feoktistov A. et.al. Mitigating uncertainty in developing and applying scientific applications in an integrated computing environment. *Programming and computer software*, 2020, 46, 8, pp. 483-502.
13. Sidorov I., Kostromin R., Feoktistov A. System for monitoring parameters of functioning infrastructure objects and their external environment. *Proceedings of the 2nd International workshop on information, computation, and control systems for distributed environments*, CEUR-WS Proceedings, 2020, v. 2638, pp. 252-264, DOI: 10.47350/ICCS-DE.2020.23.
14. Kostromin R., Feoktistov A., Voskoboinikov M. Service-oriented tools for automating digital twin development. *Proceedings of the 4th Scientific-practical workshop on information technologies: algorithms, models, systems (ITAMS 2021)*, CEUR-WS Proceedings, 2021, v. 2984, pp. 95-100, DOI: 10.47350/ITAMS.2021.12.
15. Kostromin R., Feoktistov A. Agent-based DevOps of software and hardware resources for digital twins of infrastructural objects. *Proceedings of the 4th International conference on future networks and distributed systems (ICFNDS 2020)*, ACM, 2020, pp. 1-6, DOI: 10.1145/3440749.

16. Kostromin R., Basharina O., Feoktistov A., Sidorov I. Microservice-based approach to simulating environmentally-friendly equipment of infrastructure objects taking into account meteorological data. *Atmosphere*, 2021, v. 12, no. 9, 1217. pp. 1-24, DOI: 10.3390/atmos12091217.
17. Basharina O.Yu., Feoktistov A.G. Metodika rascheta i prognozirovaniya vremennykh ryadov prirodno-klimaticheskikh pokazateley [Methods for identifying and forecasting the components of time series of natural and climatic indicators]. *Trudy mezhdunarodnoy nauchno-prakticheskoy konferentsii "Proryvnyye nauchnyye issledovaniya kak dvigatel' nauki"* [Proc. of International scientific and practical conference "Breakthrough scientific research as a science engine"], Novosibirsk, Aeterna, 2021, pp. 20-24, available at: <https://aeterna-ufa.ru/sbornik/NK-347.pdf#page=20> (accessed: 05/28/2023).
18. Edelev A.V. Otsenka trendov, regulyarnykh sezonnykh i sluchaynykh kolebaniy prirodno-klimaticheskikh kharakteristik s otsenkoy sredy R [Evaluation of trends, regular seasonal and random fluctuations in natural and climatic characteristics using the R environment]. *Trudy mezhdunarodnoy nauchno-prakticheskoy konferentsii "Sposoby, modeli i algoritmy usovershenstvovaniya nauki v sovremennykh usloviyakh"* [Proc. of International scientific and practical conference "Methods, models and algorithms for the modernization of science in modern conditions"], Ufa, Aeterna, 2020, pp. 11-14, available at: <https://aeterna-ufa.ru/sbornik/NK-318.pdf#page=11> (accessed: 05/28/2023).
19. Feoktistov A.G., Kostromin R.O., Sidorov I.A. et al. Tsifrovyye dvoyniki protsessov raboty prirodosberegayushchego oborudovaniya infrastruktornogo ob'yekta [Digital twins of operation processes for environmentally friendly equipment of infrastructure object]. *Sovremennyye naukoymkiye tekhnologii* [Modern high technologies], 2021, no. 1, pp. 57-62, DOI: 10.17513/snt.38471.
20. Karamov D.N., Edelev A.V., Feoktistov A.G. Modelirovaniye energosnabzheniya ob'yektov, okhranyayemykh Baykal'skoy prirodnoy territoriyey na osnove prirodnykh istochnikov energii [Modeling of energy supply for objects of the protected Baikal natural territory based on renewable energy sources]. *Mezhdunarodnyy tekhniko-ekonomicheskyy zhurnal* [International technical and economic journal], 2020, no. 5, pp. 7-24, DOI: 10.34286/1995-4646-2020-74-5-7-24.
21. Gorsky S., Edelev A., Feoktistov A. Data processing in problem-solving of energy system vulnerability based on in-memory data grid. *Mathematics and its applications in new computer systems (MANCS 2021). Lecture notes in networks and systems*, Springer, Cham, 2022, v. 424, pp. 271-279.
22. Edelev A.V. N.M. Beresneva, Gorsky S.A. Analiz uyazvimosti energeticheskoy infrastruktury i yego realizatsiya [Energy infrastructure vulnerability analysis and implementation]. *Sovremennyye naukoymkiye tekhnologii* [Modern high technologies], 2022, no. 1, pp. 47-52, DOI: 10.17513/snt.39008.
23. Poulin C.R., Kane M.B. Infrastructure resilience curves: Performance measures and summary metrics. *Reliability engineering & system safety*, 2021, v. 216, p. 107926, DOI: 10.1016/j.res.2021.107926.
24. Liu X., Fang Y.P., Zio E. A hierarchical resilience enhancement framework for interdependent critical infrastructures. *Reliability engineering & system safety*, 2021, v. 215, p. 107868.
25. Liu X., Ferrario E., Zio E. Identifying resilient-important elements in interdependent critical infrastructures by sensitivity analysis. *Reliability engineering & system safety*, 2019, v. 189, p. 423-434.
26. Johansson J., Hassel H. Modelling, simulation and vulnerability analysis of interdependent technical infrastructures. *Risk and interdependencies in critical infrastructures: a guideline for analysis*, London, Springer-Verlag, 2012, pp. 49-66.

Edelev Alexey Vladimirovich. *Candidate of Technical Sciences, Researcher, Melentiev Energy Systems Institute. The main direction of research – resilience of energy systems, mathematical modeling, distributed computing. AuthorID: 710350, SPIN: 8447-9522, ORCID: 0000-0003-2219-9754, flower@isem.irk.ru, Russia, Irkutsk, Lermontova, 130.*

Karamov Dmitriy Nikolaevich, *Candidate of Technical Sciences, Associate Professor, National Research Irkutsk State Technical University, Baikal School of BRICS. The main direction of research - energy systems and power engineering complexes, renewable energy sources, mathematical modelling. AuthorID: 905773, SPIN: 9977-1303, ORCID: 0000-0001-5360-4826. dmitriy.karamov@mail.ru, Russia, Irkutsk, Lermontova, 83.*

Basharina Olga Yurievna. *Candidate of Technical Sciences, Associate Professor, Researcher at the Matrosov Institute for System Dynamics and Control Theory of SB RAS, Associate Professor, Ural State University of Economics. The main direction of research - system analysis and modeling of complex systems. AuthorID: 702499, SPIN: 2612-2891, ORCID: 0000-0002-7151-782X. basharinaolga@mail.ru, Russia, Irkutsk, Lermontova, 134.*

Статья поступила в редакцию 19.11.2023; одобрена после рецензирования 23.11.2023; принята к публикации 23.11.2023.

The article was submitted 11/19/2023; approved after reviewing 11/23/2023; accepted for publication 11/23/2023.

Информационные технологии

УДК 004.415.2:004.421:005.53:61

DOI: 10.25729/ESI.2023.32.4.011

Разработка системы поддержки принятия врачебных решений на примере решения задачи выбора наилучшей траектории лечения для детей с эндокринопатиями

Берестнева Ольга Григорьевна¹, Лызин Иван Александрович¹, Аксенов Сергей Владимирович¹, Марухина Ольга Владимировна¹, Степаненко Нина Петровна²

¹Томский политехнический университет, Россия, Томск, ogbb@yandex.ru

²Федеральное государственное бюджетное учреждение «Сибирский федеральный научно-клинический центр Федерального медико-биологического агентства», Россия, Северск

Аннотация. Произведен комплексный анализ предметной области. Построены структурная схема системы поддержки принятия врачебных решений по модулям и функциональная модель, позволяющая точно отобразить задачи, выполняющие системой поддержки принятия врачебных решений. Предложен подход к предварительной обработке данных, включающий алгоритмическую базу системы поддержки принятия врачебных решений при выборе траектории лечения пациентов, анализ выбросов, обработку пропущенных данных и их восстановление.

Ключевые слова: Система поддержки принятия врачебных решений, многомерные данные, анализ, медицинские исследования, снижение размерности, классификация

Цитирование: Берестнева О.Г. Разработка системы поддержки принятия врачебных решений на примере решения задачи выбора наилучшей траектории лечения для детей с эндокринопатиями / О.Г. Берестнева, И.А. Лызин, С.В. Аксенов, О.В. Марухина, Н.П. Степаненко // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 127-138. – DOI:10.25729/ESI.2023.32.4.011.

Введение. В настоящее время врачебные ошибки представляют собой серьезную проблему при обследовании и лечении пациентов. По данным экспертов Всемирной организации здравоохранения, более 100 000 человек ежегодно умирают от врачебных ошибок. Ситуация усугубляется тем, что объемы знаний в медицине растут экспоненциально, а время на принятие врачом соответствующего решения при этом не увеличивается, что, в свою очередь, также приводит к росту числа врачебных ошибок, вследствие чего снижается эффективность работы систем здравоохранения.

Врачебная ошибка – это неблагоприятное последствие медицинской помощи, независимо от того, является ли оно очевидным или вредным для пациента. К числу проблем, которые обычно возникают при оказании медицинской помощи, относятся нежелательные реакции на лекарства, недостаточное и чрезмерное лечение, хирургические травмы и хирургические вмешательства в неправильном месте, ожоги и ошибочные диагнозы пациента. Высокая частота ошибок с серьезными последствиями чаще всего возникает в отделениях интенсивной терапии, операционных и отделениях неотложной помощи. Выделяют виды врачебных ошибок на разных стадиях:

- Диагностика (ошибка или задержка в диагностике; отказ от использования в рамках названных тестов; использование устаревших тестов или методов лечения; невыполнение действий, основанных на результатах мониторинга или тестирования).
- Лечение (ошибка при выполнении операции, процедуры или теста; ошибка в назначении лечения; ошибка в дозе или методе использования препарата; неизбежная задержка в лечении или в ответ на ненормальный тест; неподходящий уход).
- Превентивная ошибка (неспособность обеспечить профилактическое лечение; несовершенный мониторинг или последующее лечение).

Учитывая все вышеперечисленное, возникает необходимость в создании систем поддержки принятия решений. Система поддержки принятия решений (СППР) – это

интеллектуальная компьютерная программа (или комплекс программ), которая может проводить анализ, консультировать, ставить диагноз, прогнозировать процесс лечения и его результат [1]. Целью системы поддержки принятия клинических решений является улучшение оказания медицинской помощи путем усиления медицинских решений с помощью целенаправленных клинических знаний, данных о состоянии здоровья и другой информации о пациентах.

Цель исследования заключается в разработке прототипа системы поддержки принятия врачебных решений (СППВР) на основе методов интеллектуального анализа данных. В рамках данной работы будут рассмотрены следующие вопросы:

1. Обзор и анализ текущего состояния предметной области, методов и алгоритмов построения систем поддержки принятия клинических решений.
2. Разработка технологии предварительной обработки и визуализации исходных данных с последующим выявлением скрытых закономерностей и значимых предикторов.
3. Разработка алгоритмического, программного и математического обеспечения системы поддержки принятия клинических решений.

Обзор и анализ текущего состояния предметной области. В настоящее время многие учреждения здравоохранения разного масштаба и профилей во всем мире широко применяют автоматизацию медицинских технологий и соответствующих бизнес-процессов. Для этого используются медицинские информационные системы. Одним из важных направлений является разработка СППР, которые используют для диагностики и прогнозирования заболеваний. Подобные системы не могут нести ответственность за принятые с ее помощью решения, но способны существенно упростить и ускорить работу врача. Цель СППР заключается в осуществлении кооперации системы и человека в процессе принятия решений.

Исследования в области систем поддержки принятия клинических решений начались с работ Ледли Р.С. (1959) [2], Барнетт Г.О. (1968) [3], Китова А.И. (1971) [1], Ван Беммель Дж.Х. (1984) [4], Осипова Г.С. (2005) [5], и активно продолжаются по сей день, с применением методов машинного обучения при разработке таких систем (Лебедев Г.С. [6], Назаренко Г.И. [5], Осипов Е. [7, 8], Манн и др. [9], Гарвин Дж. Х. и др. [10], Блекер С. и др. [11], Сильвейра Д. В. и др. [12], Ковальчук С. В. и др. [13]).

В большинстве публикаций, посвященных медицинским информационным системам, обычно выделяют два основных типа: электронная медицинская карта (ЭМК) и поддержка принятия клинических решений (CDS – clinical decision support). В отечественных стандартах [14] медицинские информационные системы подразделяются на пять видов, среди которых аналогом CDS являются «медико-технологические информационные системы, предназначенные для обеспечения процессов диагностики, лечения, реабилитации и профилактики больных в учреждениях здравоохранения, функциональное назначение которых трактуется достаточно широко» [15]. В функции CDS входит своевременное напоминание врачам, информирование о необходимых анализах, аллергии на лекарства и т.д.

Среди ученых, проводивших исследования в области интеллектуальных систем, можно назвать такие имена, как Гаврилова Т.А., Финн В.К., Хорошевский В.Ф., Борисов А.Н., Мамдани Э., Заде Л., Вагин Б.Н., Стефанюк В.Л., Ларичев О.И., Осипов Г.С., Аверкин А.Н., Фоминых И.Б., Попов Е.В., Поспелов Д.А., Васильев С.Н., Ясный Дж. и др.

Среди отечественных разработок можно назвать системы ЭКОНА, АССОД, ОТЭКС (Новосибирск). Особо следует выделить инструментальный комплекс IMSCLOG (Янковская А.Е., Томск), Универсальный классификатор Юдина В.Ш. (Москва) и систему Deep Data Diver Дюка В.А. и Асеева М.Г. (Санкт-Петербург).

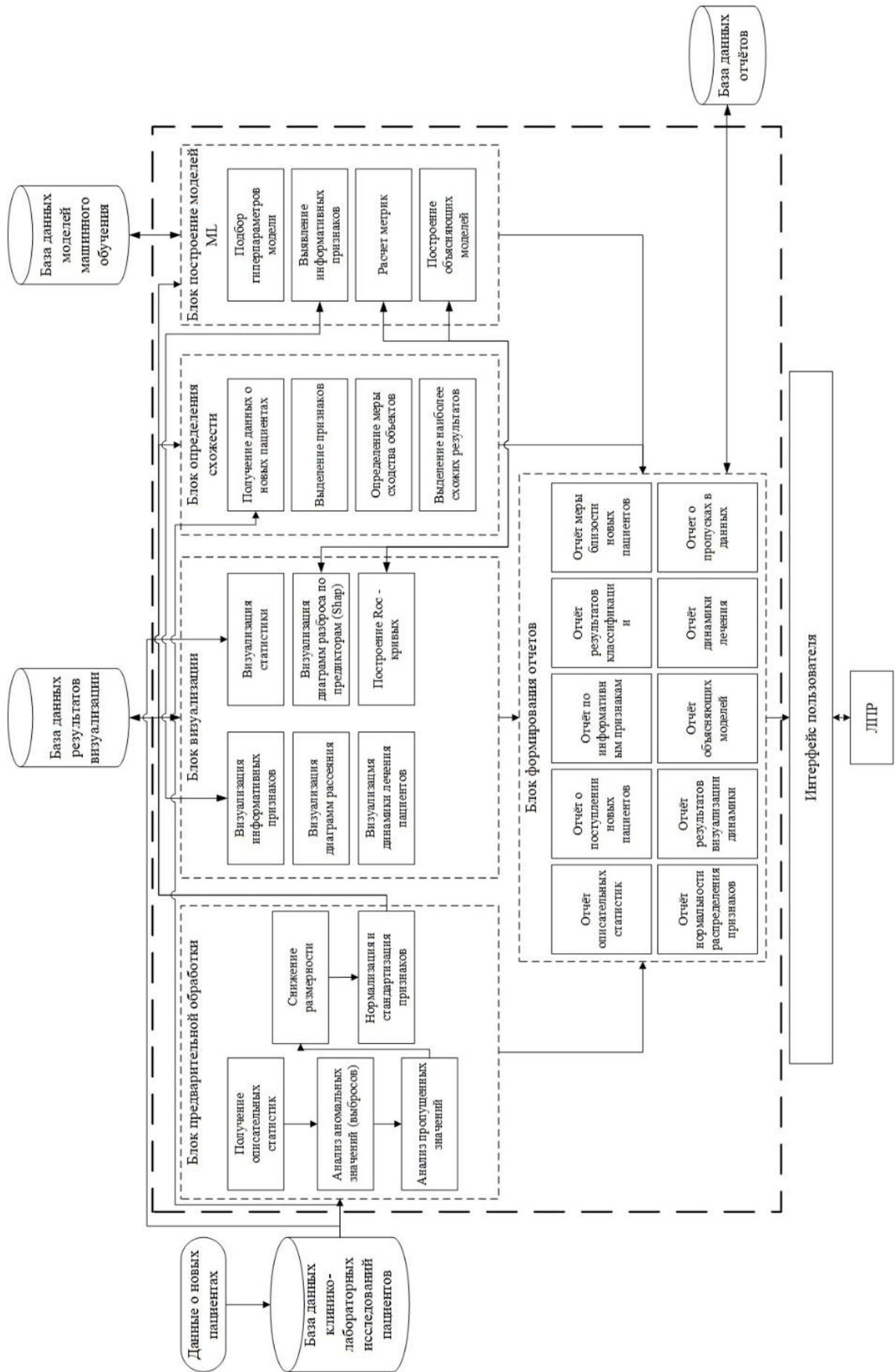


Рис. 1. Схема системы поддержки принятия решений

Среди зарубежных программ широко используется система WIZWHY (WizSoft, США), основанная на алгоритмах ограниченного перебора для поиска логических закономерностей в данных. С 2000-го года успешно развиваются системы: IndiGO (формирование индивидуальных протоколов диагностики и лечения); Auminence (формирование диагностического плана); DiagnosisOne (поиск ошибок и формирование планов лечения), Isabel Healthcare (анализ симптомов); VisualDx (дифференциальная диагностика); Nuance (СППР для радиологии); IBM Watson (суперкомпьютер IBM с вопросно-ответной системой искусственного интеллекта). В направлении СППВР развиваются и более простые программы для конечного пользователя (WebMD Symptom Checker, DrNow, iPharmacy, EasyDiagnosis). Большинство ведущих СППВР систем имеют в настоящее время мобильные и онлайн-версии.

В России в 80-90 гг. XX века были разработаны экспертные медицинские системы: ДИН (диагностика неотложных состояний); «Айболит» (диагностика острых расстройств кровообращения); ДИАГЕН (диагностика наследственных заболеваний); SYNGEN (хромосомная патология) и ряд других. Эти системы существенно устарели и не получили дальнейшего развития, на смену им приходят другие, с более совершенными IT-инструментами, алгоритмами, методами, средствами, а также используемыми медицинскими знаниями.

Исходя из проведенного анализа предметной области была спроектирована структурная схема СППВР, которая представлена на рисунке 1.

Построение функциональной модели поможет точно отобразить, какие задачи должна решать данная система. Укрупненная функциональная модель системы представлена на рисунке 2.

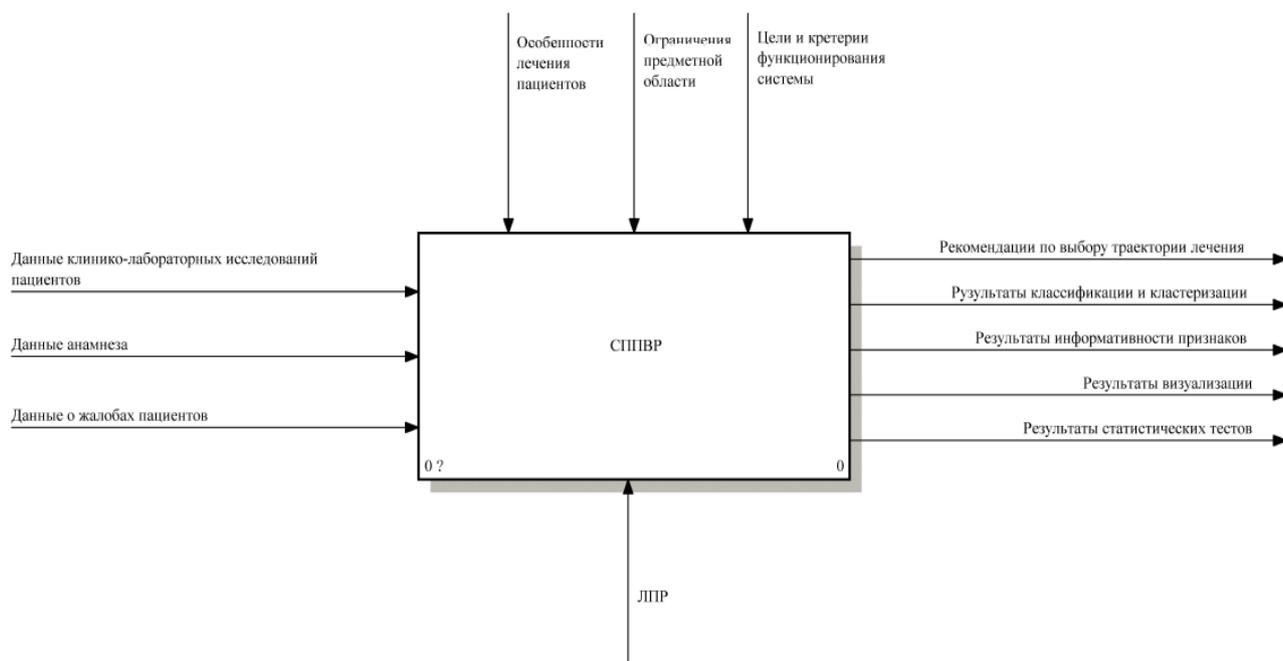


Рис. 2. Функциональная модель СППВР

Декомпозиция контекстной диаграммы IDEF0 (рис. 2) представлена на рисунке 3.

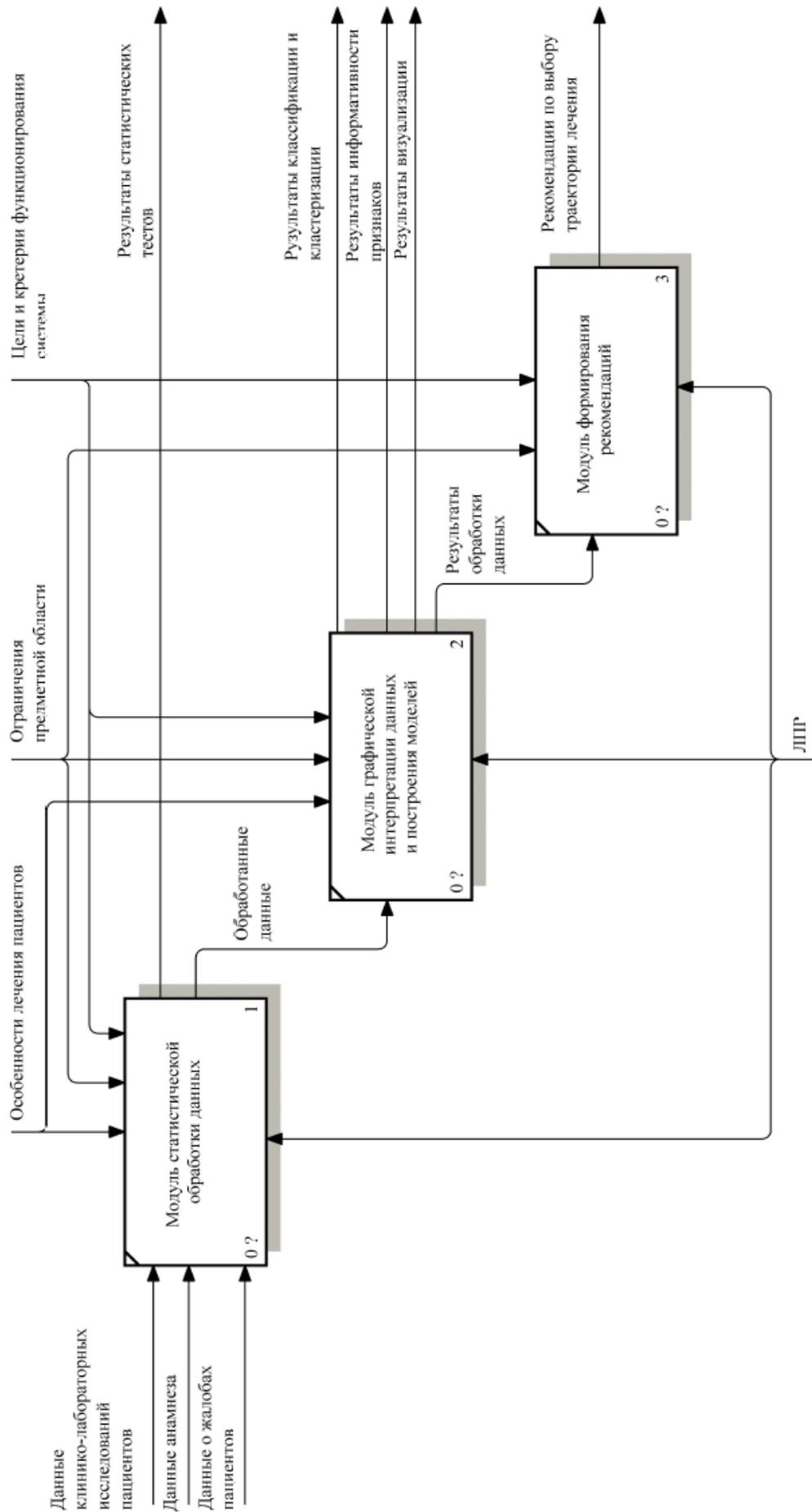


Рис. 3. Декомпозиция функциональной модели СПДВР

Разработка технологии предварительной обработки. В рамках сотрудничества с Томским научно-исследовательским институтом курортологии и физиотерапии (ТНИИКиФ) был получен набор данных клинического исследования, проведенного на базе детского отделения в период с 2006 по 2013 год, в котором приняли участие 464 ребенка, страдающих ожирением в возрасте от 7 до 18 лет. Данные представлены в виде многомерной таблицы и содержат значения групп клинико-лабораторных показателей пациентов. Набор данных содержит 345 объектов (пациентов), 160 показателей (значения до и после лечения) и 5 групп (траекторий лечения).

Все показатели относятся к разным физиологическим группам: 1) клиника, 2) сердечно-сосудистая система; 3) физическая работоспособность; 4) липидный обмен; 5) биохимия крови; 6) углеводный обмен; 7) гормональный статус; 8) цитокиновый статус; 9) иммунологический статус; 10) состояние калликреин-кининовой системы; 11) окислительная способность плазмы крови. Помимо перечисленных показателей, набор данных содержит информацию о принадлежности пациента к группе лечения (1-5). Разрабатываемая система предназначена для поддержки принятия врачебного решения о назначении пациенту наиболее подходящей для него траектории лечения на основе анализа его первичных клинико-лабораторных показателей.

Предварительная подготовка данных. Первичная обработка данных направлена на упорядочивание информации об исследуемом объекте. Она может быть направлена на упорядочивание исходных данных; обнаружение и устранение ошибок, некорректно введенных данных, выбросов, пробелов в сведениях; выявление скрытых закономерностей и связей. Обработка данных при проведении исследования осуществлялась по алгоритму, схема которого представлена на рисунке 4.

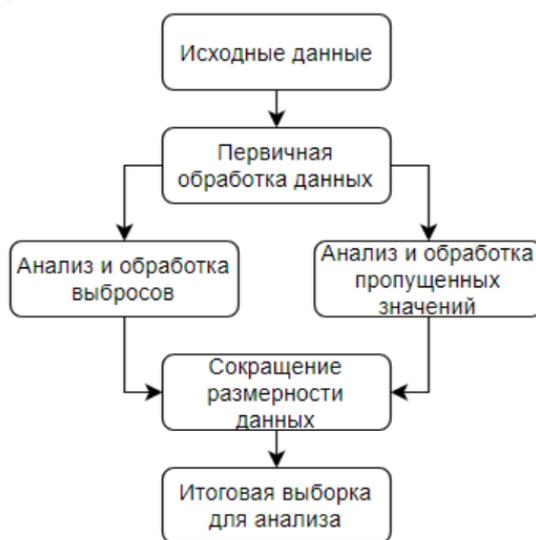


Рис. 4. Схема алгоритма первичной обработки данных

Задачей исследования является определение траектории лечения новых пациентов, поэтому в рабочей выборке содержатся данные о пациентах по клинико-лабораторным показателям до лечения и назначенные им методики лечения (группы).

Важным этапом подготовки данных к восстановлению является анализ выбросов, который позволяет избежать отклонения восстановленных данных от разумных значений. Для определения выбросов воспользуемся графическим анализом наблюдений посредством диаграмм размахов, с помощью которых легко распознать значения, лежащие за пределами наблюдений. На рисунке 5 представлена диаграмма размаха, реализованная в R, для одного из показателей.

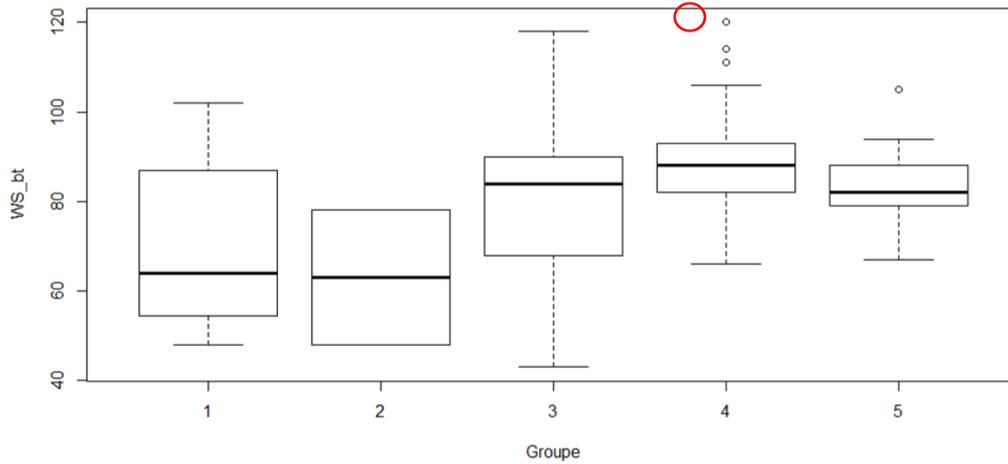


Рис. 5. Диаграмма размаха для показателя WS_bt (объем талии до лечения)

Выбросы наблюдаются в четвертой и пятой группах. В данном случае работа с выбросами – задача нетривиальная и решение об исключении того или иного выброса принимается только коллегиально с экспертом-врачом. Таким образом, было принято решение об исключении из выборки только одного наблюдения, указанного на рисунке 5 красным кружком. Аналогично были проанализированы различные случаи наличия и обработки выбросов по всем показателям во всех группах лечения.

Данные исходного набора очень разрежены, в частности, некоторые признаки являются полностью пустыми на всем промежутке записей. Подобные признаки были исключены из рабочей выборки. Также были исключены все признаки до лечения и разности между значений до и после лечения. На диаграмме (рис. 6) слева представлено количество пропущенных значений для каждого показателя в отдельности, а справа – для комбинации показателей.

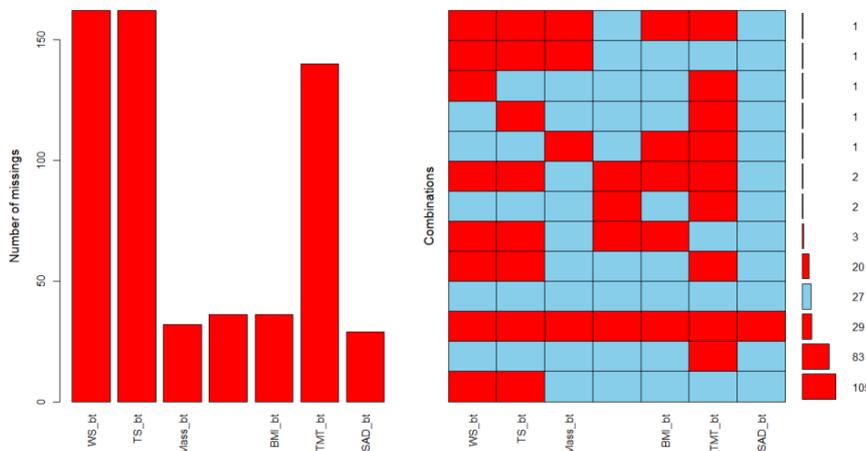


Рис. 6. Фрагмент диаграммы пропущенных значений в исходном наборе данных

Для восстановления пропущенных значений использовался метод множественного восстановления, при котором заполнение пропусков происходит при помощи повторного моделирования. Из существующего набора данных с пропущенными значениями создаются несколько полных наборов данных, к каждому из которых применяются стандартные статистические методы для формирования окончательных результатов. Идея множественного восстановления пропущенных данных хорошо реализована в таком пакете R, как *miss*. В *miss* многомерное восстановление данных реализуется при помощи связанных уравнений. Пропущенные значения замещаются при помощи выборок Гиббса. По умолчанию значения каждой переменной, содержащей пропущенные значения, предсказываются по значениям остальных переменных. Полученные уравнения используются для замещения пропущенных

данных подходящими значениями. Этот процесс повторяется, пока значения для пропущенных данных не сойдутся

Для оценки точности восстановленных данных были построены графики плотности распределения исходных (показана красным цветом) и заполненных данных (показана синим цветом), пример представлен на рисунок 7. Исходя из этого, можно сделать вывод что распределения восстановленных данных близко к исходным.

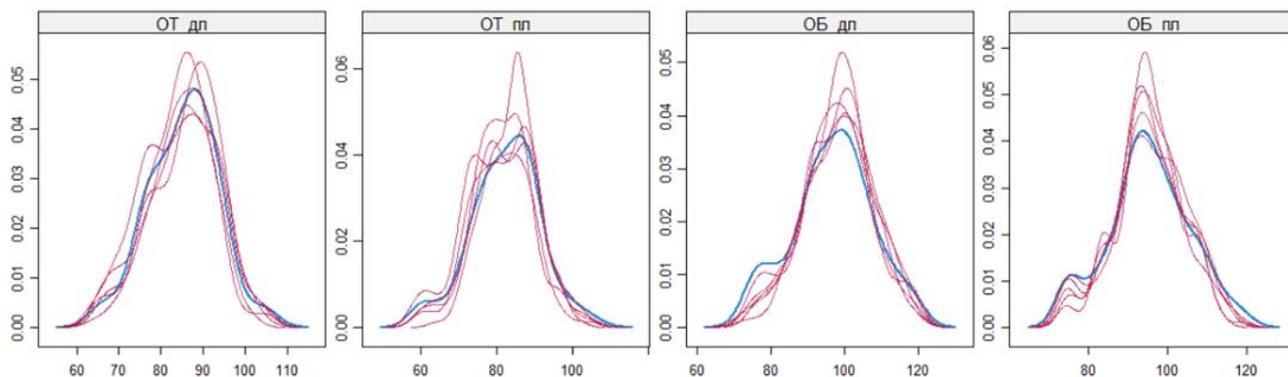


Рис. 7. Графики плотности распределения исходных (синий) и заполненных (красный) данных

Сокращение размерности пространства признаков. Для сокращения размерности существует два типа методов (рисунок 8).



Рис. 8. Классификация методов уменьшения размерности

Методы сокращения размерности с преобразованием переменных, как правило, сопровождаются проблемой интерпретации новых переменных и определения вклада исходных переменных, поэтому при решении поставленных задач данные методы использоваться не будут.

Рассмотрим методы выбора параметров для уменьшения размерности исходного набора данных. Высокая корреляция между двумя переменными означает, что они имеют сходные тенденции и, вероятно, несут схожую информацию. Это может резко снизить производительность некоторых моделей. Возможно рассчитать корреляцию между независимыми числовыми переменными; если коэффициент корреляции пересекает определенное пороговое значение, можно отбросить одну из переменных. Отбрасывание переменной очень субъективно и всегда должно выполняться с учетом предметной области. Как правило, предлагается сохранять те переменные, которые показывают среднюю или высокую корреляцию с целевой переменной.

Важность переменных может быть оценена из данных путем построения модели. Случайный лес (Random Forest) является одним из наиболее широко используемых алгоритмов выбора объектов. В него уже встроена функция важности, поэтому нет необходимости рассчитывать ее отдельно. На рисунке 9 представлены значения важности для атрибутов набора данных:

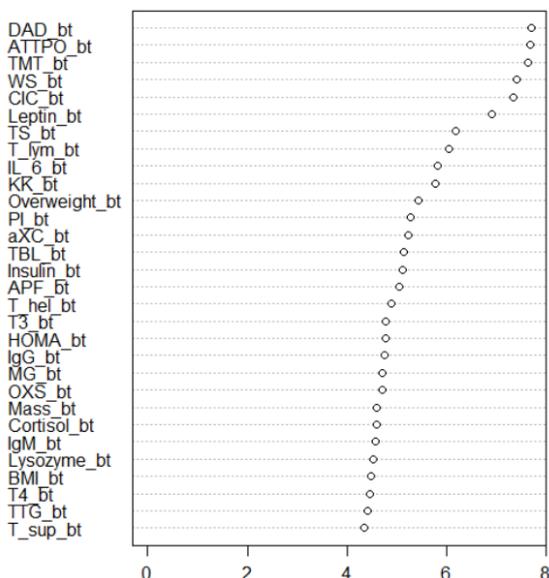


Рис 9. Фрагмент графика градации признаков по степени важности (Random Forest)

После предварительной обработки сокращения размерности, было принято решение оставить 35 информативных атрибутов для дальнейшего исследования.

Классификация данных. В ходе исследования были построены несколько различных моделей классификации, которые давали различную точность. В качестве примера приведем результат классификации с использованием деревьев решений на полученном после предварительной обработки наборе данных. Реализация построения и графическое представление дерева решения в R выглядят следующим образом (рисунок 10).

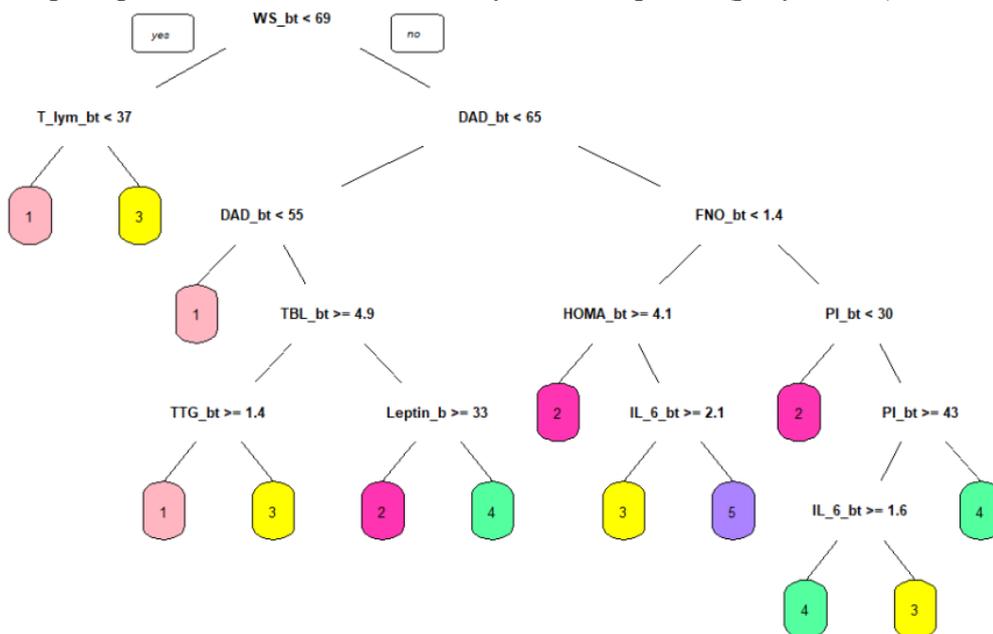


Рис 10. Графическое построение дерева решений в R

Точность данной модели на тестовых данных 0.8284906 – это одна из лучших моделей из серии экспериментов. В дальнейшем планируется повышение точности за счет увеличения объема исходных данных и мониторинга качества их сбора.

В результате проведенного исследования был разработан алгоритм поддержки принятия решения о выборе траектории лечения детей с избыточным весом на основе построения дерева решения. Для разработки алгоритма были проведены следующие этапы:

- Получены результаты клинического исследования, проведенного на базе детского отделения Томского НИИ курортологии и физиотерапии.
- Проведена первичная обработка данных с помощью методов языка R. Был проведен анализ выбросов, восстановление пропущенных значений методом множественного восстановления, а также сокращение размерности за счет выявления информативных признаков.
- Предложена модель классификации пациентов по группам лечения, выбрана модель с наилучшим значениям точности.

Заключение. В рамках выполнения исследования были проведены обзор и анализ текущего состояния предметной области, методов и алгоритмов построения систем поддержки принятия клинических решений. Спроектирована структурная схема системы поддержки принятия врачебных решений и схема в нотации IDEF0. Разработан алгоритм технологии предварительной обработки и визуализации исходных данных с последующей классификацией по заданным группам.

Структурная схема СППВР и алгоритм предварительной обработки и визуализации исходных данных несомненно могут быть использованы при создании любых экспертных СППВР.

Благодарности. Исследование выполнено при частичной финансовой поддержке РФФИ в рамках выполнения научных проектов № 19-37-90005.

Список источников

1. Китов А.И. Основные принципы построения ИПС для медицины / А.И. Китов // Цифровая вычислительная техника и программирование. – М.: Сов. радио, 1971. – Вып. 6. – С. 17-31.
2. Ledley R.S. Digital electronic computers in biomedical science. *Science*, 1959, vol. 130, pp. 1225-1234.
3. Barnett G.O. Computers in patient care. *N. Engl. J. Med*, 1968, vol. 279, no. 24, pp. 1321-1327.
4. van Bemmel J.H. The structure of medical informatics. *Med. Inform (Lond)*, 1984 Jul-Dec, vol. 9, no. 3-4, pp. 175-80.
5. Назаренко Г.И. Медицинские информационные системы: теория и практика / Г.И. Назаренко, Я.И. Гулиев, Д.Е. Ермаков. – М.: Физматлит, 2005. – 320 с.
6. Лебедев Г.С. Классификация медицинских информационных систем / Г.С. Лебедева, Ю.Ю. Мухина // Информационные технологии в медицине. – М.: Радиотехника, 2012. – С. 42-62.
7. Kleyko D., Osipov E., Wiklund U. Hyperdimensional Computing framework for analysis of cardiorespiratory synchronization during paced deep breathing. *IEEE Access PP (99)*, 2019, vol. 7, pp. 34403-34415, DOI: 10.1109/ACCESS.2019.2904311.
8. Wedekind D., Kleyko D., Osipov E., et al. Robust methods for automated selection of cardiac signals after blind source separation. *IEEE transactions on biomedical engineering*, 2018, vol. 65(10), pp. 2248-2258, DOI: 10.1109/TBME.2017.2788701.
9. Mann D., Hess R., McGinn T., et al. Adaptive design of a clinical decision support tool: What the impact on utilization rates means for future CDS research. *Digit Health*, 2019, vol. 5, pp.1-12.
10. Garvin J.H., Ducom J., Matheny M., et al. Descriptive usability study of CirrODS: Clinical decision and workflow support tool for management of patients with cirrhosis. *JMIR Med Inform*, 2019, vol. 7(3): e13627.
11. Blecker S., Pandya R., Stork S., et al. Interruptive versus noninterruptive clinical decision support: usability study. *JMIR Hum Factors*, 2019, vol. 6(2): e12469.
12. Silveira D.V., Marcolino M.S., Machado E.L. Development and evaluation of a mobile decision support system for hypertension management in the primary care setting in brazil: mixed-methods field study on usability, feasibility, and utility. *JMIR Mhealth Uhealth*. 2019, vol. 7(3):e9869.
13. Информационные системы в здравоохранении. Общие требования. СТО МОСЗ 91500.16.0002-2004
14. Ваганова Е.В. Медицинские информационные системы как объект оценки: факторы и тенденции развития. / Е.В. Ваганова // Вестник Томского государственного университета. Экономика, 2017. – № 37. – С. 113-130.

15. Berestneva O.G., Lyzin I.A., Stepanenko N.P. Application of integral health index in evaluation of efficiency of rehabilitation of children with obesity. *Studies in systems, decision and control*, 2022, vol. 416, pp. 87-100. DOI:10.1007/978-3-030-95112-2_8.

Лызин Иван Александрович. Аспирант Отделения информационных технологий ИШИТР ФГАОУ ВО Национального исследовательского Томского политехнического университета (ФГАОУ ВО НИ ТПУ), ORCID: 0000-0003-2827-441X, AuthorID: 1155694, Lyzin@tpu.ru, 634050, Россия, г. Томск-50, проспект Ленина, дом 30,

Берестнева Ольга Григорьевна. Доктор технических наук, профессор Отделения информационных технологий ИШИТР ФГАОУ ВО Национального исследовательского Томского политехнического университета (ФГАОУ ВО НИ ТПУ), ORCID: 0000-0002-4243-0637; SPIN: 8026-4116, eLibrary AuthorID: 112998, ogb6@yandex.ru, 634050, Россия, г. Томск-50, проспект Ленина, дом 30.

Аксенов Сергей Владимирович. Кандидат технических наук, доцент Отделение информационных технологий ИШИТР Томского политехнического университета, axuonov@tpu.ru. г. Томск, просп. Ленина, 30. 634034.

Марухина Ольга Владимировна. Кандидат технических наук, доцент Отделение информационных технологий ИШИТР Томского политехнического университета, marukhina@tpu.ru. г. Томск, просп. Ленина, 30. 634034.

Степаненко Нина Петровна. Кандидат медицинских наук, педиатр высшей категории, Сибирский федеральный научно-клинический центр Федерального медико-биологического агентства. ORCID 0000-0001-6844-9208, stepanekonina62@bk.ru.

UDC 004.415.2:004.421:005.53:61

DOI: 10.25729/ESI.2023.32.4.011

Development of a medical decision support system on the example of solving the problem of choosing the best treatment trajectory for children with endocrinopathies

**Olga G. Berestneva¹, Ivan A. Lyzin¹, Sergey V. Aksenov¹, Olga V. Marukhina¹,
Nina P. Stepanenko²**

¹National Research Tomsk Polytechnic University, Russia, Tomsk, ogb6@yandex.ru

²Siberian Federal Scientific and Clinical Center of Federal Medical and Biological Agency, Russia, Seversk, stepanekonina62@bk.ru

Abstract. A comprehensive analysis of the subject area was made. A block diagram of the medical decision support system for modules has been built, and a functional model that allows you to accurately display the tasks performed by the medical decision support system. An approach to data preprocessing is proposed, including the algorithmic base of a medical decision support system when choosing a trajectory for treating patients, analysis of outliers, processing of missing data and their recovery.

Keywords: medical decision support system, multidimensional data, analysis, medical research, dimensionality reduction, classification

Acknowledgements: the study was carried out with the partial financial support of the Russian Foundation for Basic Research (RFBR) as part of research projects No. 19-37-90005

References

1. Kitov A.I. Osnovnye principy postroeniya IPS dlya mediciny [Basic principles of building IPS for medicine]. Tsifrovaya vychislitel'naya tekhnika i programmirovaniye [Digital computing and programming], Moscow: Sov. radio, 1971, iss. 6, pp. 17-31.
2. Ledley R.S. Digital electronic computers in biomedical science. *Science*, 1959, vol. 130, pp. 1225-1234.
3. Barnett G.O. Computers in patient care. *N. Engl. J. Med.*, 1968, vol. 279, no. 24, pp. 1321-1327.

4. van Bommel J.H. The structure of medical informatics. *Med. Inform (Lond)*, 1984 Jul-Dec, vol. 9, no. 3-4, pp. 175-80.
5. Nazarenko G.I., Guliev YA.I., Ermakov D.E. *Medicinskie informacionnyye sistemy: teoriya i praktika* [Medical information systems: theory and practice]. M.: Fizmatlit, 2005. 320 p.
6. Lebedev G.S., Mukhina Yu.Yu. *Klassifikaciya medicinskih informacionnyh sistem* [Classification of medical information systems]. *Informacionnyye tekhnologii v meditsine*, M.: Radiotekhnika [Information technologies in medicine, Moscow: Radio Engineering], 2012. pp. 42-62.
7. Kleyko D., Osipov E., Wiklund U. Hyperdimensional Computing framework for analysis of cardiorespiratory synchronization during paced deep breathing. *IEEE Access PP (99)*, 2019, vol. 7, pp. 34403-34415, DOI: 10.1109/ACCESS.2019.2904311.
8. Wedekind D., Kleyko D., Osipov E., et al. Robust methods for automated selection of cardiac signals after blind source separation. *IEEE transactions on biomedical engineering*, 2018, vol. 65(10), pp. 2248-2258, DOI: 10.1109/TBME.2017.2788701.
9. Mann D., Hess R., McGinn T., et al. Adaptive design of a clinical decision support tool: What the impact on utilization rates means for future CDS research. *Digit Health*, 2019, vol. 5, pp.1-12.
10. Garvin J.H., Ducom J., Matheny M., et al. Descriptive usability study of CirrODS: Clinical decision and workflow support tool for management of patients with cirrhosis. *JMIR Med Inform*, 2019, vol. 7(3): e13627.
11. Blecker S., Pandya R., Stork S., et al. Interruptive versus noninterruptive clinical decision support: usability study. *JMIR Hum Factors*, 2019, vol. 6(2): e12469.
12. Silveira D.V., Marcolino M.S., Machado E.L. Development and evaluation of a mobile decision support system for hypertension management in the primary care setting in brazil: mixed-methods field study on usability, feasibility, and utility. *JMIR Mhealth Uhealth*. 2019, vol. 7(3):e9869.
13. *Informacionnyye sistemy v zdravoohranenii* [Information systems in healthcare]. General requirements. STO MOSZ 91500.16.0002-2004.
14. Vaganova E.V. *Medicinskie informacionnyye sistemy kak ob"ekt ocenki: faktory i tendencii razvitiya*. [Medical information systems as an object of evaluation: factors and development trends]. *Vestnik Tomskogo gosudarstvennogo universiteta. Ekonomika* [Bulletin of Tomsk state university. Economy], 2017, no. 37, pp. 113-130.
15. Berestneva O.G., Lyzin I.A., Stepanenko N.P. Application of integral health index in evaluation of efficiency of rehabilitation of children with obesity. *Studies in systems, decision and control*, 2022, vol. 416, pp. 87-100, DOI:10.1007/978-3-030-95112-2_8.

Berestneva Olga Grigorievna. *Dr. Sci. (Tech.), Professor, Department of Information Technologies, School of Computer Science & Robotics, National Research Tomsk Polytechnic University. ORCID 0000-0002-4243-0637; SPIN: 8026-4116, AuthorID: 112998, ogb6@yandex.ru, 634050, Russia, Tomsk-50, prospect Lenina, 30,*

Lyzin Ivan Aleksandrovich. *Postgraduate Student, Department of Information Technologies, School of Computer Science & Robotics, National Research Tomsk Polytechnic University. ORCID 0000-0003-2827-441X, AuthorID: 1155694, Lyzin@tpu.ru, 634050, Russia, Tomsk-50, prospect Lenina, 30.*

Aksenov Sergey Vladimirovich. *Candidate of Engineering Science, assistant professor, Department of information technology, School of Engineering Information Technology and Robotics, Tomsk Polytechnic University (TPU). axyonov@tpu.ru*

Marukhina Olga Vladimirovna. *Candidate of Engineering Science, assistant professor, Department of information technology, School of Engineering Information Technology and Robotics, Tomsk Polytechnic University (TPU). ORCID 0000-0003-1834-9692. marukhina@tpu.ru*

Stepanenko Nina Petrovna. *Cand. Sci. (Med.), Pediatrician of Superior Expert Category, Siberian Federal Scientific and Clinical Center of Federal Medical and Biological Agency. ORCID 0000-0001-6844-9208.: stepanenkoina62@bk.ru.*

Статья поступила в редакцию 18.05.2023; одобрена после рецензирования 16.10.2023; принята к публикации 16.12.2023.

The article was submitted 05/18/2023; approved after reviewing 10/16/2023; accepted for publication 12/16/2023.

УДК 528.92

DOI: 10.25729/ESI.2023.32.4.012

Применение FABDEM и других современных цифровых моделей рельефа в системе аграрного мониторинга

Кузнецова Анна Сергеевна¹, Пушкарев Александр Александрович¹, Краснощекон Константин Вячеславович¹, Якубайлик Олег Эдуардович², Ерунова Марина Геннадьевна¹

¹Федеральный исследовательский центр «Красноярский научный центр СО РАН», Россия, Красноярск, anna.simakina.97@mail.ru

²Институт вычислительного моделирования СО РАН, Россия, Красноярск

Аннотация. В работе рассматривается получение морфометрических характеристик рельефа с применением современных методов дистанционного зондирования Земли (ДЗЗ) и геоинформационных технологий для обеспечения системы аграрного мониторинга ФИЦ КНЦ СО РАН. В качестве исходных данных использованы новая цифровая модель рельефа (ЦМР) глобального масштаба FABDEM и высокоточные аэрофотоснимки, полученные при помощи съемки с беспилотного летательного аппарата (БПЛА) и специализированного геодезического оборудования. На их основе в геоинформационных системах (ГИС) QGIS, ArcGIS и Спутник Агро разработан ряд тематических карт, отображающих основные морфометрические характеристики рельефа. При составлении карт соблюдены требования, предъявляемые к картографическим материалам (выбор правильной картографической проекции, формата, типовой легенды и т.д.). Проведено сравнение морфометрических характеристик рельефа, подготовленных по набору данных FABDEM и съемке с БПЛА. Высокая степень сходства между ними демонстрирует, что ЦМР FABDEM является подходящим вариантом для получения информации о рельефе местности при отсутствии возможности проведения съемки с БПЛА. В результате получен большой объем информации об особенностях устройства рельефа поверхности опытно-производственного хозяйства (ОПХ) «Михайловское». Для хранения, распространения и анализа сведений об устройстве поверхности исследуемого объекта тематические карты, сохраненные в формате проектов qgis, импортированы в систему аграрного мониторинга ФИЦ КНЦ СО РАН. Совместное использование тематических карт о рельефе местности и данных, уже входящих в систему (спутниковые данные, вегетационные индексы, почвенные и климатические характеристики), позволит учесть состояние сельскохозяйственных угодий при разработке стратегий по управлению и использованию земель сельскохозяйственного назначения.

Ключевые слова: агромониторинг, информационные системы, FABDEM, БПЛА, картографирование

Цитирование: Кузнецова А.С. Применение FABDEM и других современных цифровых моделей рельефа в системе аграрного мониторинга / А.С. Кузнецова, А.А. Пушкарев, К.В. Краснощекон, О.Э. Якубайлик, М.Г. Ерунова // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 139-147. – DOI: 10.25729/ESI.2023.32.4.012.

Введение. Сельское хозяйство является одной из важнейших отраслей экономики в мире и играет ключевую роль в обеспечении продовольственной безопасности. За последние два десятилетия агропроизводство столкнулось с рядом проблем, таких, как изменение климата, экстремальные погодные условия, растущий спрос на продукты питания и т. д. [1, 2]. Для решения текущих проблем в сельскохозяйственный сектор необходимо внедрять новые технологии, которые позволят отслеживать происходящие изменения.

В настоящее время для эффективного управления землями сельскохозяйственного назначения активно разрабатываются и внедряются специализированные информационные системы аграрного мониторинга [3-5]. В большинстве случаев существующие системы направлены на сбор, хранение, обработку, распространение и анализ данных ДЗЗ и информационных продуктов, получаемых на их основе. Спутниковые данные используют для мониторинга посевных и уборочных работ, а также для других агротехнических мероприятий, выполняемых на полях [6]. Информация о вегетационных индексах и метеорологических условиях позволяет проводить оценку состояния сельскохозяйственных посевов и степени

влияния совокупности факторов на урожайность [7, 8]. Однако для точного анализа на уровне сельскохозяйственных угодий в информационных системах аграрного мониторинга необходимо учитывать особенности рельефа местности [9].

Рельеф местности влияет на интенсивность освещения и прогревания поверхности, термический и ветровой режимы, распределение осадков, протекание эрозионных процессов, почвообразование и другие аспекты [10, 11]. Влияние рельефа описывается морфометрическими характеристиками, такими, как уклон, крутизна и форма склонов, экспозиция относительно сторон света. От них зависит выбор сельскохозяйственных культур, технологии обработки почв, системы орошения и дренажа.

Современные достижения в области геоинформационных технологий и ДЗЗ способны обеспечить потребность сельского хозяйства в картографических сведениях о рельефе местности. Для получения морфометрических характеристик сельскохозяйственных угодий высокого пространственного разрешения используют БПЛА [12]. В глобальном масштабе для изучения рельефа требуется ЦМР, получаемая при помощи космических измерений. К распространенным ЦМР глобального масштаба в открытом доступе относятся продукты MERIT DEM (с разрешением ~90 м), SRTM (с разрешением ~30 м) и другие [13]. Наиболее новой и точной ЦМР глобального масштаба считается FABDEM (Forest And Buildings removed Copernicus DEM). Это первая глобальная ЦМР, в которой удалены высоты деревьев и зданий [14].

Целью данной работы является получение морфометрических характеристик рельефа с использованием современных методов ДЗЗ и геоинформационных технологий для обеспечения системы аграрного мониторинга ФИЦ КНЦ СО РАН [15].

1. Объект, материалы и методы исследования. В качестве объекта исследования выбрано ОПХ «Михайловское» филиала ФГБНУ ФИЦ КНЦ СО РАН. Территория хозяйства расположена в северо-западной части Ужурского района, в 30 км от районного центра г. Ужур, 250 км от г. Красноярск. Общая площадь ОПХ «Михайловское» составляет 15209 га, в том числе площадь сельскохозяйственных угодий – 11294 га.

В начале работы была подготовлена цифровая карта сельскохозяйственных угодий ОПХ «Михайловское» в ГИС QGIS на основе высококачественного космического снимка. В качестве базовой проекции используется система координат WGS 84 / UTM zone 46N. Цифровая карта угодий (полей) состоит из 51 объекта, каждый объект содержит информацию о номере поля, севообороте и площади (га).

Аэрофотосъемка территории ОПХ проведена при помощи камеры Sony RX1, установленной на борту F1xar 007 в ходе полевых исследований в 2022 году. Съемка осуществлялась с высоты 300 м. Пространственное разрешение полученных снимков – 4 см/пиксель. Для увеличения точности координат центров аэрофотоснимков дополнительно использовались геодезический RTK приемник Emlid Reach M2, установленный на борту БПЛА, и ГНСС приёмник Emlid Reach RS2, стационарно расположенный на земле. В программе TOPOSETTER полученные с приемников данные сопоставлялись со снимками Sony RX1. Это способствовало увеличению точности координат центров каждого кадра с 3-5 м (при использовании стандартных GPS) до 7-10 см.

Массив изображений, полученный в ходе аэрофотосъемки с БПЛА, был обработан в программном обеспечении Agisoft Metashape Professional. Изображения камеры Sony RX1 проходят два этапа предварительной обработки: выравнивание и оптимизацию. Далее для каждого снимка строятся карты глубины и создается плотное облако точек. По плотному облаку точек проводится классификация точек рельефа. Точки «земли» отделяются от других точек (растительности, зданий, опор ЛЭП и др.). Результатами обработки исходных снимков

являются ортофотоплан местности и ЦМР с пространственным разрешением 4 см и 15 см соответственно.

В качестве исходных данных среднего разрешения для получения морфометрических характеристик рельефа на территории исследования использовалась ЦМР FABDEM. ЦМР разработана Бристольским университетом и доступна по всему миру с шагом сетки в 1 угловую секунду (приблизительно 30 м на экваторе) [14].

По ЦМР FABDEM среднего разрешения, ортофотоплану местности и ЦМР высокого разрешения выполнялось картографирование морфометрических характеристик рельефа в ГИС QGIS, ArcGIS и Спутник Агро.

Технология разработки тематических карт по набору данных FABDEM в ГИС QGIS состоит из нескольких этапов: перепроецирование исходного слоя FABDEM и его обрезка по границам хозяйства; использование инструментов морфометрического анализа «создать изолинии», «экспозиция» и «крутизна»; вычисление преобладающей экспозиции полей при помощи модуля «зональная статистика».

Вычисление тематической карты индекса Topographic Wetness Index (TWI) проведено в ГИС ArcGIS [16]. В первую очередь были подготовлены топографические слои: удельная площадь водосбора участка и угол наклона с использованием инструментов «flow direction», «flow accumulation» и «slope».

В ГИС Спутник Агро подготовлены тематические карты по ЦМР и ортофотоплану местности высокого пространственного разрешения при помощи инструментов программы «моделирование водных стоков» и «расчет уклонов».

2. Результаты. На основе набора данных FABDEM были получены следующие тематические карты:

- ЦМР и тематическая карта горизонталей (с расстоянием между горизонталями 10 м) демонстрируют устройство поверхности территории и позволяют определять характер перепада высот. Такого рода информация необходима для планирования размещения сельскохозяйственных угодий, прогнозирования развития эрозионных процессов и их воздействия на сельскохозяйственные угодья.
- Тематическая карта экспозиции содержит данные о рассчитанных показателях преобладающей экспозиции для всех сельскохозяйственных угодий ОПХ «Михайловское». Карта классифицирована по восьми румбам экспозиции для определения полей с более благоприятным световым и термическим режимом.
- Тематическая карта уклонов служит для визуализации уклонов местности, разграниченных по группам, с целью определения участков на местности с высокими уклонами. Информация о значении уклона, его расположении и протяженности крайне важна при определении назначения сельскохозяйственных угодий и проведении технических операций (укос сенокосов, сбор урожая, подкормка).
- Завершающим результатом построения карт по набору данных FABDEM является тематическая карта индекса TWI. Индекс TWI представляет собой устойчивый индекс влажности, учитывающий следующие ключевые компоненты: общую площадь водосбора и уклон склона. Обычно используется для определения гидрологических путей потока и всестороннего учета влияния характеристик местности на распределение влаги.

Высокое пространственное разрешение и точность привязки аэрофотоснимков способствовали построению ЦМР и тематической карты уклонов в границах сельскохозяйственных угодий с более высокой точностью и детальностью. Кроме этого, по данным БПЛА подготовлены тематические карты поверхностного стока и накопления воды,

содержащие линейные и площадные объекты соответственно. Визуализация сети водотоков и участков, на которых задерживается вода, помогает понять направление водотоков и определить переувлажненные участки на сельскохозяйственных угодьях.

Тематические карты, отображающие основные морфометрические характеристики рельефа ОПХ «Михайловское», разработанные в различных ГИС, были приведены к единым параметрам и системе координат (WGS 84 / UTM zone 46N) в программе QGIS. Исходные слои преобразованы в форматы GeoTIFF и shape. Каждому из них подобрана типовая легенда. Готовые тематические карты сохранены в качестве проектов qgs формата. Для хранения, распространения и анализа сведений об устройстве рельефа поверхности исследуемого объекта полученные тематические карты импортированы в систему аграрного мониторинга.

Информационная система аграрного мониторинга ФИЦ КНЦ СО РАН предназначена для мониторинга земель сельскохозяйственного назначения Красноярского края. Сочетает в себе удобный и понятный интерфейс для пользователей. Направлена на распространение данных компьютерных моделей (информация о климатических характеристиках), спутниковых снимков и продуктов, получаемых на их основе (вегетационные индексы), тематических карт (цифровые карты полей, агрохимические характеристики почвы и др.).

Импорт тематических карт в систему осуществляется через веб-форму создания нового слоя на странице администрирования. Форма предусматривает заполнение основных полей (название, описание, доступ, локация, тип данных) и загрузку архива с подготовленным проектом. Последующие этапы загрузки проектов полностью автоматизированы. После отправки формы проект проверяется на наличие ошибок и помещается в специализированное хранилище. Данное хранилище находится в области видимости QGIS сервера, который, в свою очередь, отвечает за визуализацию тематических карт в системе аграрного мониторинга.

Благодаря тематическим картам, построенным на основе набора данных FABDEM и съемки с БПЛА, получен большой объем информации об особенностях рельефа территории хозяйства «Михайловское» в целом. Высоты поверхности исследуемой территории колеблются в пределах от 295 м до 500 м (рис. 1а). Большинство сельскохозяйственных угодий расположено на северных, южных и юго-восточных склонах (рис. 1б). Также были определены сельскохозяйственные угодья, где находятся крупные участки с высокими значениями уклонов, и выявлено большое количество гидрологических путей потока по индексу TWI.

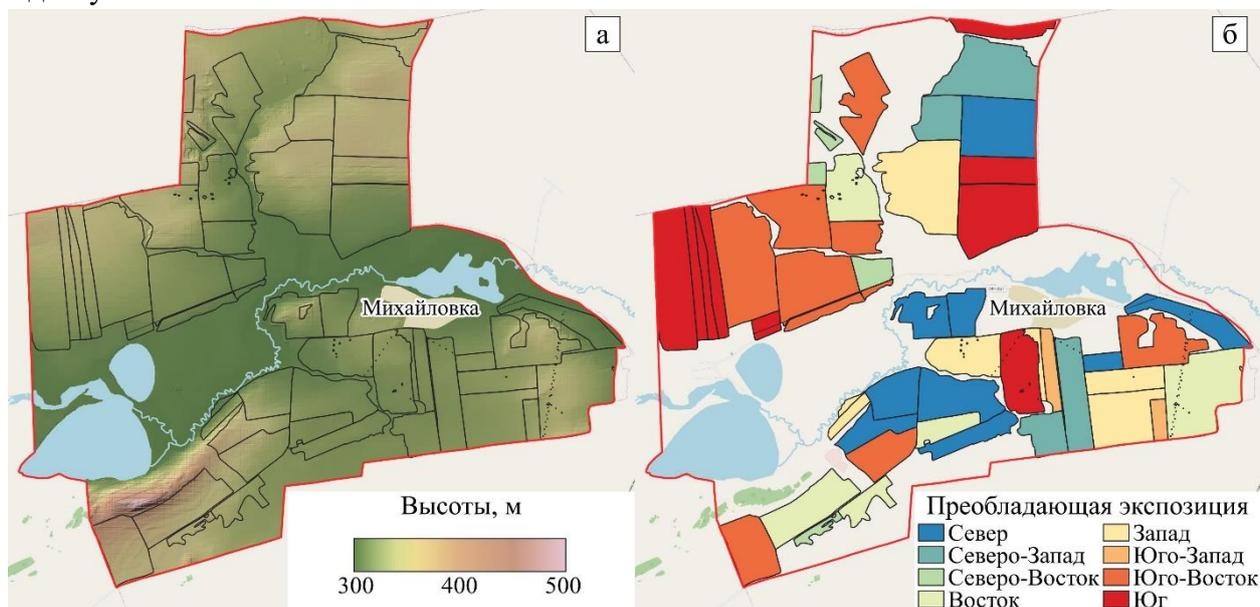


Рис. 1. Тематические карты по набору данных FABDEM: а – карта высот; б – карта преобладающей экспозиции полей

Проведение полевых работ БПЛА корректирует использование ЦМР глобального масштаба. Так, например, по данным FABDEM определена группа полей ОПХ «Михайловское» с высокими значениями уклонов (рис. 2а). Построение тематической карты по данным БПЛА (рис. 2б) позволило уточнить значения уклонов и их расположение, а также выявить новые области с высокими значениями уклонов.

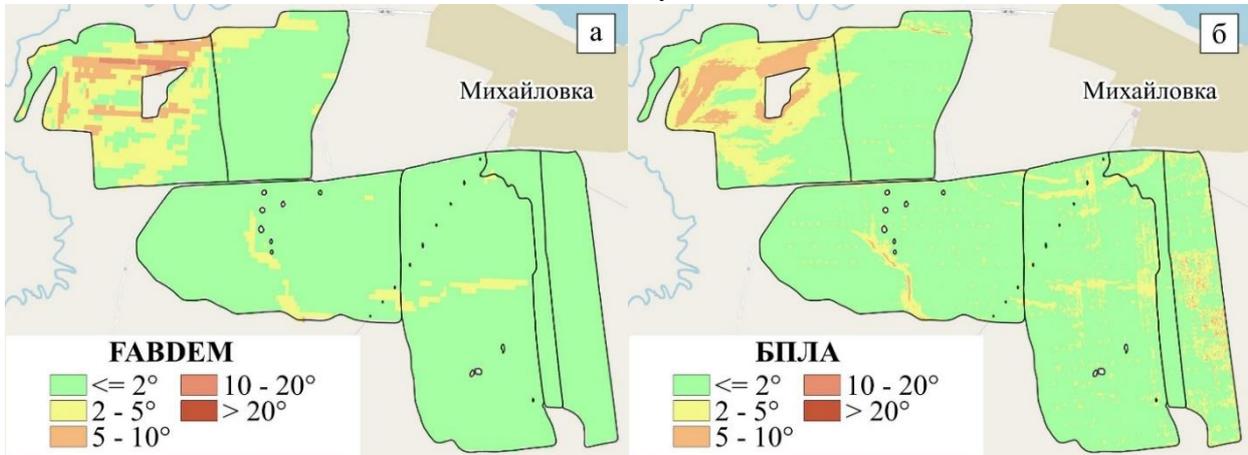


Рис. 2. Фрагменты тематических карт уклонов на примере группы полей ОПХ «Михайловское»

В качестве еще одного примера приведены тематические карты индекса TWI (рис. 3а) и поверхностного стока воды (рис. 3б), которые предоставляют сведения о гидрологических путях потока воды. Тематическая карта индекса TWI подходит для определения путей потоков в границах полей на крупных территориях. При этом тематическая карта, построенная по данным БПЛА, демонстрирует пути потоков воды и их ответвления с более высокой точностью.

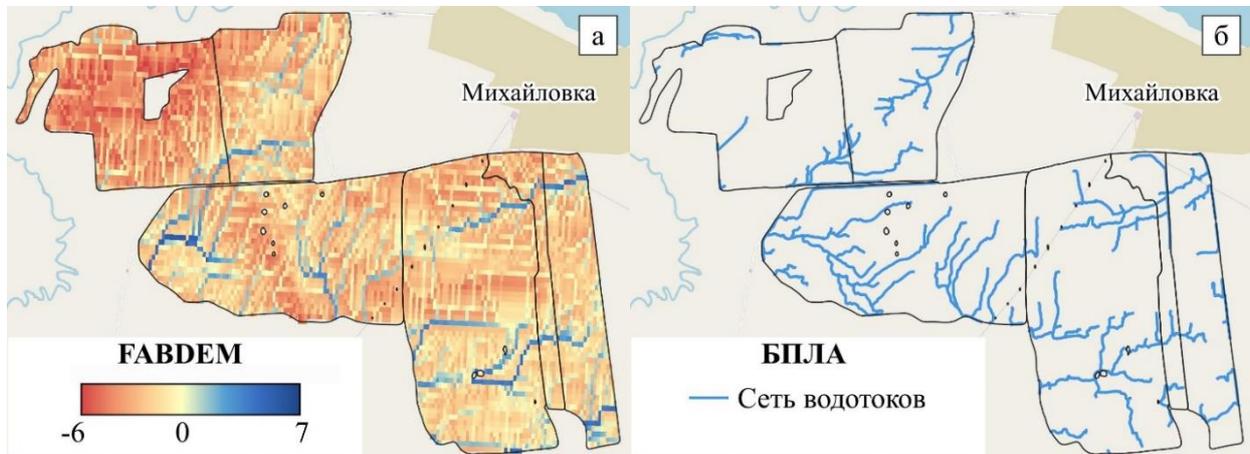


Рис. 3. Фрагменты тематических карт с распределением путей потоков воды: а – индекс TWI; б – карта поверхностного стока воды

Высокое разрешение аэрофотоснимков позволяет строить высокоточные тематические карты, характеризующие особенности рельефа местности. К сожалению, не для каждой территории сельскохозяйственного производства возможно проведение съемки с БПЛА. Это связано с различными географическими, экономическими, социальными и другими факторами.

Сравнение морфометрических характеристик рельефа, подготовленных по набору данных FABDEM и съемке с БПЛА, демонстрирует высокую степень сходства между ними. Это указывает на то, что новая и точная ЦМР FABDEM является подходящим вариантом для

получения информации о рельефе местности при отсутствии возможности проведения съемки с БПЛА. При этом точечное проведение работ по сбору информации при помощи БПЛА отдельных участков повысит качество проведения аграрного мониторинга.

Заключение. В результате выполненных работ при помощи современных методов ДЗЗ и геоинформационных технологий получен ряд тематических карт, отображающих основные морфометрические характеристики рельефа ОПХ «Михайловское». Для их подготовки использованы новая и точная ЦМР глобального масштаба FABDEM и аэрофотоснимки с высокой точностью координат центров. В будущем планируется получить морфометрические характеристики рельефа и на другие сельскохозяйственные территории Красноярского края, включив их в информационную систему аграрного мониторинга ФИЦ КНЦ СО РАН.

Обеспечение системы аграрного мониторинга морфометрическими характеристиками рельефа предоставит доступ специалистам аграрного сектора к полноценной информации об особенностях рельефа местности. Совместное использование карт о рельефе местности и данных, уже входящих в систему, позволит учесть состояние сельскохозяйственных угодий при разработке стратегий по управлению и использованию земель сельскохозяйственного назначения.

Список источников

1. Wu L., Elshorbagy A., Helgason W. Assessment of agricultural adaptations to climate change from a water-energy-food nexus perspective, *Agricultural Water Management*, 2023, v. 284, 108343.
2. Blickensdörfer L., Schwieder M., Pflugmacher D., et al. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany, *Remote Sensing of Environment*, 2022, v.269, 112831.
3. Fritz S., See L., Laso Bayas J.C., et al. A comparison of global agricultural monitoring systems and current gaps, *Agricultural Systems*, 2019, v.168, pp. 258-272.
4. Денисов П.В. Возможности и опыт использования информационной системы Вега-PRO для мониторинга сельскохозяйственных земель / П.В. Денисов, К.А. Трошко, Е.А. Лупян и др. // Вычислительные технологии. Электрон. журн., 2022. – Т. 27. – № 3. – С. 66-83. – DOI: 10.25743/ICT.2022.27.3.006.
5. Zahran S.A.E.-S., Saeed R.A.-H., Elazizy I.M. Remote sensing based water resources and agriculture spatial indicators system, *The Egyptian Journal of Remote Sensing and Space Science*, 2022, v.25, pp. 515-527.
6. Гаврилов В.К. Возможности спутникового мониторинга для оценки состояния сельскохозяйственных земель и их инвестиционной привлекательности / В.К. Гаврилов, М.В. Зимин, Н.А. Моисеева // Управление рисками в АПК. Электрон. журн., 2021. – № 1(39). – С. 8-21. – DOI: 10.53988/24136573-2021-01-01.
7. Плотников Д.Е. Развитие метода автоматического распознавания озимых культур на основе спутниковых данных для оценки их состояния на территории Республики Крым / Д.Е. Плотников, Е.С. Ёлкина, Е.А. Дунаева [и др.] // Таврический вестник аграрной науки. Электрон. журн., 2020. – № 1(21). – С. 64-82. – DOI: 10.33952/2542-0720-2020-1-21-64-83.
8. Степанов А.С. Влияние климатических характеристик и значений вегетационного индекса NDVI на урожайность сои (на примере районов Приморского края) / А.С. Степанов, Т.А. Асеева, К.Н. Дубровин // Аграрный вестник Урала. Электрон. журн., 2020. – № 1(192). – С. 10-19. – DOI: 10.32417/1997-4868-2020-192-1-10-19.
9. Rebelo C., Nascimento J. Measurement of soil tillage using UAV high-resolution 3D Data. *Remote Sensing*, 2021, v.13, no.21, 4336.
10. Плисенко О.А. Расчет основных геоморфологических параметров элементарных поверхностей рельефа на основе ЦРМ в цифровой платформе точного земледелия Адыгеи / О.А. Плисенко, Т.П. Варшанина // Вестник Адыгейского государственного университета. Серия 4: Естественно-математические и технические науки. Электрон. журн., 2021. – № 4(291). – С. 125-134. – DOI: 10.53598/2410-3225-2021-4-291-125-134.
11. Мажитова Г.З. Совершенствование методики крупномасштабного агроландшафтного картографирования на основе применения беспилотных летательных аппаратов / Г.З. Мажитова, С.В. Пашков, С.В. Крыцкий // Региональные геосистемы. Электрон. журн., 2020. – Т. 44. – № 1. – С. 64-74. – DOI: 10.18413/2712-7443-2020-44-1-64-74.
12. Пашков С.В. Применение ГИС-технологий и аэрофотосъемки для геоинформационного картографирования и моделирования рельефа агроландшафтов / С.В. Пашков, Г.З. Мажитова // Известия

Иркутского государственного университета. Серия: Науки о Земле. Электрон. журн., 2020. – Т. 34. – С. 82-95. – DOI: 10.26516/2073-3402.2020.34.82.

13. Uuemaa E., Ahi S., Montibeller B., et al. Vertical accuracy of freely available global digital elevation models (ASTER, AW3D30, MERIT, TanDEM-X, SRTM, and NASADEM), Remote Sensing, 2020, v.12, no.21, 3482.
14. Hawker L., Uhe P., Paulo L., et al. A 30 m global map of elevation with forests and buildings removed, Environmental Research Letters, 2022, v.17, no. 2, 024016.
15. Pushkarev A.A., Yakubailik O.E. A web application for visualization, analysis, and processing of agricultural monitoring spatial-temporal data. CEUR Workshop Proceedings, Novosibirsk, 2021, pp. 231-237.
16. Копецký М., Мачек М., Wild J. Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition, Science of the total environment, 2021, v.757, 143785.

Кузнецова Анна Сергеевна. Младший научный сотрудник Федерального исследовательского центра «Красноярский научный центр СО РАН». AuthorID: 1123552, SPIN: 4915-3961, ORCID: 0000-0002-8351-7536, anna.simakina.97@mail.ru, Россия, г. Красноярск, Академгородок, д. 50.

Пушкарев Александр Александрович. Младший научный сотрудник Федерального исследовательского центра «Красноярский научный центр СО РАН». AuthorID: 1019386, SPIN: 5344-5762, ORCID: 0009-0000-5790-966X, flash550@yandex.ru, Россия, г. Красноярск, Академгородок, д. 50.

Краснощеков Константин Вячеславович. Младший научный сотрудник Федерального исследовательского центра «Красноярский научный центр СО РАН». AuthorID: 1005418, SPIN: 2937-3719, ORCID: 0000-0001-9477-4127, krasko@icm.krasn.ru, Россия, г. Красноярск, Академгородок, д. 50.

Якубайлик Олег Эдуардович. Кандидат физико-математических наук, зам. директора по научной работе Института вычислительного моделирования СО РАН. AuthorID: 99904, SPIN: 7820-4178, ORCID: 0000-0002-2668-4776, oleg@icm.krasn.ru, Россия, г. Красноярск, Академгородок, д. 50, стр. 44.

Ерунова Марина Геннадьевна. Кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Красноярский научный центр СО РАН». AuthorID: 116643, SPIN: 5424-1229, ORCID: 0000-0002-1922-1036, marina.erunova@gmail.com, Россия, г. Красноярск, Академгородок, д. 50.

UDC 528.92

DOI: 10.25729/ESI.2023.32.4.012

Application of FABDEM and other modern digital elevation models in the agricultural monitoring system

Anna S. Kuznetsova¹, Alexander A. Pushkarev¹, Konstantin V. Krasnoshchekov¹, Oleg E. Yakubailik², Marina G. Erunova¹

¹Federal Research Center Krasnoyarsk Science Center of the SB RAS, Russia, Krasnoyarsk, anna.simakina.97@mail.ru

²Institute of Computational Modelling SB RAS, Russia, Krasnoyarsk

Abstract. The paper considers obtaining morphometric characteristics of the relief using modern methods of remote sensing of the Earth and geoinformation technologies to provide the agricultural monitoring system of the FRC KSC SB RAS. The new digital terrain model (DEM) of the global scale FABDEM and high-precision aerial photographs obtained by shooting from an unmanned aerial vehicle (UAV) and specialized geodetic equipment were used as initial data. Based on them, a number of thematic maps have been developed in the geoinformation systems QGIS, ArcGIS and Sputnik Agro, displaying the main morphometric characteristics of the relief. When making maps, the requirements for cartographic materials are met (choosing the correct cartographic projection, format, standard legend, etc.). The morphometric characteristics of the relief prepared using the FABDEM data set and UAV imagery were compared. The high degree of similarity between them demonstrates that the FABDEM DEM is a suitable option for obtaining information about the terrain in the absence of the possibility of shooting from a UAV. As a result, a large amount of information was obtained about the features of the device of the surface relief of the Mikhailovskoye agricultural experimental production facility (EPF). To store, distribute and analyze information about the structure of the surface of the object under study, thematic maps saved in the qgis project format were imported into the agricultural monitoring system of the FRC KSC SB RAS. The joint use of thematic

maps on the terrain and data already included in the system (satellite data, vegetation indices, soil and climatic characteristics) will allow taking into account the state of agricultural land when developing strategies for the management and use of agricultural land.

Keywords: agromonitoring, information systems, FABDEM, UAV, mapping

References

1. Wu L., Elshorbagy A., Helgason W. Assessment of agricultural adaptations to climate change from a water-energy-food nexus perspective, *Agricultural Water Management*, 2023, v.284, 108343.
2. Blickensdörfer L., Schwieder M., Pflugmacher D., et al. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany, *Remote Sensing of Environment*, 2022, v.269, 112831.
3. Fritz S., See L., Laso Bayas J.C., et al. A comparison of global agricultural monitoring systems and current gaps, *Agricultural Systems*, 2019, v.168, pp. 258-272.
4. Denisov P.V., Troshko K.A., Lupjan E.A. et al. Vozmozhnosti i opyt ispol'zovaniya informacionnoj sistemy Vega-PRO dlya monitoringa sel'skoxozyajstvennyx zemel [Potential and experience of Vega-PRO information system use for monitoring of agricultural lands]. *Computational Technologies*, 2022, vol.27, no. 3, pp. 66-83, DOI: 10.25743/ICT.2022.27.3.006.
5. Zahran S.A.E.-S., Saeed R.A.-H., Elazizy I.M. Remote sensing based water resources and agriculture spatial indicators system, *The Egyptian Journal of Remote Sensing and Space Science*, 2022, v.25, pp. 515-527.
6. Gavrilov V.K., Zimin M.V., Moiseeva N.A. Vozmozhnosti sputnikovogo monitoringa dlya ocenki sostoyaniya sel'skoxozyajstvennyx zemel i ix investicionnoj privlekatel'nosti [Satellite monitoring capabilities for assessing the state of agricultural land and its investment attractiveness]. *Upravlenie riskami v APK [Risk management in the AIC]*, 2021, no.1(39), pp. 8-21, DOI: 10.53988/24136573-2021-01-01.
7. Plotnikov D.E., Jolkina E.S., Dunaeva E.A. [et al.] Razvitie metoda avtomaticheskogo raspoznavaniya ozimy x kul'tur na osnove sputnikovyx dannyx dlya ocenki ix sostoyaniya na territorii Respubliki Krym [Development of the method for automatic winter crops mapping by means of remote sensing aimed at crops state assessment over the Republic of Crimea]. *Tavricheskiy vestnik agrarnoj nauki [Tauride Bulletin of Agrarian science]*, 2020, no.1(21), pp.64-82, DOI: 10.33952/2542-0720-2020-1-21-64-83.
8. Stepanov A.S., Aseeva T.A., Dubrovin K.N. Vliyanie klimaticheskix xarakteristik i znachenij vegetacionnogo indeksa NDVI na urozhajnost' soi (na primere rajonov Primorskogo kraja) [The influence of climatic characteristics and values of NDVI at soybean yield (on the example of the districts of the Primorskiy region)]. *Agrarian Bulletin of the Urals*, 2020, no. 1(192), pp. 10-19, DOI: 10.32417/1997-4868-2020-192-1-10-19.
9. Rebelo C., Nascimento J. Measurement of soil tillage using UAV high-resolution 3D Data. *Remote Sensing*, 2021, 13(21), 4336.
10. Plisenko O.A., Varshanina T.P. Raschet osnovnyx geomorfologicheskix parametrov elementarnyx poverxnostej rel'efa na osnove CzRM v cifrovoj platforme tochnogo zemledeliya Adygei [Digital relief model based calculation of basic geomorphological parameters of elementary relief surfaces in digital platform of Adygea precision agriculture]. *Vestnik Adygejskogo gosudarstvennogo universiteta. Seriya 4: Estestvenno-matematicheskie i tehnicheckie nauki [Bulletin of the Adygea state university. Series 4: Natural-mathematical and technical sciences]*, 2021, no.4(291), pp. 125-134, DOI: 10.53598/2410-3225-2021-4-291-125-134.
11. Mazhitova G.Z., Pashkov S.V., Kryckij S.V. Sovershenstvovanie metodiki krupnomasshtabnogo agrolandshaftnogo kartografirovaniya na osnove primeneniya bespilotnyx letatel'nyx apparatov [Improvement of the methodology of large-scale agrarian landscape mapping based on uav application]. *Regional'nyye geosistemy [Regional Geosystems]*, 2020, v.44, no.1, pp.64-74, DOI: 10.18413/2712-7443-2020-44-1-64-74.
12. Pashkov S.V., Mazhitova G.Z. Primenenie GIS-tekhnologij i aerofotos'emki dlya geoinformacionnogo kartografirovaniya i modelirovaniya rel'efa agrolandshaftov [Application of GIS technologies and aerial photography for geoinformation mapping and modelling of relief of agroland landscapes]. *Izvestiya Irkutskogo gosudarstvennogo universiteta. Seriya: Nauki o Zemle [The bulletin of Irkutsk state university. Series Earth sciences]*, 2020, v.34, pp.82-95, DOI: 10.26516/2073-3402.2020.34.82.
13. Uemaa E., Ahi S., Montibeller B., et al. Vertical accuracy of freely available global digital elevation models (ASTER, AW3D30, MERIT, TanDEM-X, SRTM, and NASADEM), *Remote Sensing*, 2020, v.12, no.21, 3482.
14. Hawker L., Uhe P., Paulo L., et al. A 30 m global map of elevation with forests and buildings removed, *Environmental Research Letters*, 2022, v.17, no.2, 024016.
15. Pushkarev A.A., Yakubailik O.E. A web application for visualization, analysis, and processing of agricultural monitoring spatial-temporal data. *CEUR Workshop Proceedings, Novosibirsk*, 2021, pp. 231-237.
16. Kopecký M., Macek M., Wild J. Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition, *Science of The Total Environment*, 2021, v.757, 143785.

Kuznetsova Anna Sergeevna. Junior Researcher, Federal Research Center Krasnoyarsk Science Center SB RAS. AuthorID: 1123552, SPIN: 4915-3961, ORCID: 0000-0002-8351-7536, anna.simakina.97@mail.ru, Russia, Krasnoyarsk, Akademgorodok, 50.

Pushkarev Alexander Aleksandrovich. Junior Researcher, Federal Research Center Krasnoyarsk Science Center SB RAS. AuthorID: 1019386, SPIN: 5344-5762, ORCID: 0009-0000-5790-966X, flash550@yandex.ru, Russia, Krasnoyarsk, Akademgorodok, 50.

Krasnoshchekov Konstantin Viacheslavovich. Junior Researcher, Federal Research Center Krasnoyarsk Science Center SB RAS. AuthorID: 1005418, SPIN: 2937-3719, ORCID: 0000-0001-9477-4127, krasko@icm.krasn.ru, Russia, Krasnoyarsk, Akademgorodok, 50.

Yakubailik Oleg Eduardovich. Candidate of Physico-mathematical Sciences, Deputy Director for Scientific Work, Institute of Computational Modeling SB RAS. AuthorID: 99904, SPIN: 7820-4178, ORCID: 0000-0002-2668-4776, oleg@icm.krasn.ru, Russia, Krasnoyarsk, Akademgorodok, 50, building 44.

Erunova Marina Gennadiyevna. Candidate of Technical Sciences, Senior Researcher Federal Research Center Krasnoyarsk Science Center SB RAS. AuthorID: 116643, SPIN: 5424-1229, ORCID: 0000-0002-1922-1036, marina.erunova@gmail.com, Russia, Krasnoyarsk, Akademgorodok, 50.

Статья поступила в редакцию 15.08.2023; одобрена после рецензирования 29.11.2023; принята к публикации 16.12.2023.

The article was submitted 08/15/2023; approved after reviewing 11/29/2023; accepted for publication 12/16/2023.

УДК 005.94

DOI: 10.25729/ESI.2023.32.4.013

Разработка и апробация методики картирования знаний наукоемкой организации

Гринберг Эльвира Яковлевна¹, Верзин Евгений Андреевич²

¹ Санкт-Петербургский Университет, Россия, Санкт-Петербург,
st057454@student.spbu.ru

² ООО «Газпромнефть НТЦ», Россия, Санкт-Петербург

Аннотация. Систематическое картирование, категоризация и сравнительный анализ знаний организации позволяют принимать информированные управленческие решения. Проблемой в формировании системы интеллектуальной поддержки принятия решений на основе картирования знаний является разрозненность и неоднородность источников знаний, а также отсутствие унификации их оценки. В качестве решения описанной проблемы в данной работе представлены разработка и апробация методики картирования знаний. Отличительной особенностью представленного подхода является методологическая опора на онтологический инжиниринг. Онтология была разработана и предложена для актуализации рабочей группе. В ней были учтены следующие направления: функциональные знания, углубленные профильные знания, владение специализированным ПО, обладание знаниями о контрагентах, знания, полученные в результате работы над проектами. Сбор данных для цифровой карты осуществлялся с помощью опроса сотрудников и дополнялся информацией, предоставленной подразделением. В результате картирования созданы переиспользуемые системы организации знаний. Итогом апробации методики стал набор рекомендаций по стратегическому позиционированию картируемого подразделения.

Ключевые слова: Карта знаний, онтология, картирование компетенций, дашборд знаний

Цитирование: Гринберг Э.Я. Разработка и апробация методики картирования знаний наукоемкой организации / Э.Я. Гринберг, Е.А. Верзин // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 148-159. – DOI:10.25729/ESI.2023.32.4.013.

Введение. Организация может получить большую отдачу от усилий по стратегическому развитию, сгруппировав познавательный опыт в «критическую научающую массу» вокруг некоторых стратегических областей знания [1]. В силу индивидуальности организаций каждая из них обладает обширным набором экспертных знаний, однако представление о распределении этих знаний и о документах, недоступных другим сотрудникам, остается неполным. Результатом нехватки информации о существующих знаниевых активах является многократное решение одних и тех же задач. Знания могут быть доступны, но сотрудник не знает об их существовании и выполняет работы с нуля.

Проектная и исследовательская деятельность в нефтегазовой отрасли генерирует значительное количество новых знаниевых активов и требует их переиспользования и применения на ежедневной основе [2]. Учитывая данный фактор, взаимозависимость функциональных направлений отрасли [3], а также возможность сокращения издержек за счет тиражирования лучшего опыта Центра по управлению знаниями Научно-Технического Центра «Газпром нефти» инициировала разработку информационного решения. Его задача – установить расположение требуемых формализованных знаний либо их носителя (эксперта предметной области) для решения проектных и исследовательских задач, а также выявить направления, которые требуют особого внимания.

В качестве подхода к решению поставленной задачи была разработана и апробирована методика картирования знаний. Карты знаний создаются для решения ряда организационных задач и отображают бизнес-процессы, области знаний, роли, компетенции или стратегические потребности бизнеса [4]. Некоторые карты знаний фокусируются на экспертных должностях, тогда как многие организации концентрируются на стратегических потребностях внутри организации.

Было выбрано картирование знаний на основе процессов, подобно тому, как это ранее было реализовано для специалистов МЧС [5] и в авиастроительной отрасли [6], поскольку оно оптимально работает в организациях с четко определенными бизнес-процессами. Отличительной особенностью представленного подхода является методологическая опора на онтологический инжиниринг, в частности на [7]. Онтологии позволяют организовать знания в виде концептуальных схем, где определены типы сущностей и их связи. Такая система картирования знаний облегчает поиск, понимание и использование информации в различных областях [8]. Организации используют картирование знаний на основе ролей для представления знаний ключевых экспертов предметной области, в том числе, тех, которые выходят на пенсию. Карта знаний содержит подробное описание знаний, которые можно обобщить и формализовать, но не знания, как таковые.

В международной практике [9] выделяют три основные цели построения цифровой карты знаний:

1. Выявлять, сохранять и воспроизводить ключевые знания и экспертизу.
2. Планировать стратегические улучшения деятельности по управлению знаниями.
3. Стандартизировать и системно применять схожие практики.

Инициатором картирования, как правило, является функциональный руководитель с просьбой провести картирование знаний конкретного эксперта, команды или подразделения. Исходя из того, что руководители отвечают за стратегическое планирование и кадровое обеспечение своих функциональных направлений, они могут точно определить потенциальные пробелы в компетенциях, которые можно проанализировать с помощью картирования. Картирование, как правило, применяется для выявления и документирования знаний следующих категорий сотрудников: руководители высшего звена, профильные эксперты, владельцы процессов, другой опытный и высококвалифицированный персонал.

Специфика разработки карты знаний, которая описана ниже, заключается в сочетании бизнес-подхода и онтологического инжиниринга. Знаниевые активы могут храниться как в документах, процессах и системах, так и в головах экспертов. В данной работе в качестве источников знания рассматриваются только сотрудники, их активность в создании документов и участии в проектах.

Каждая карта – это одно визуальное представление графа знаний. Существует множество определений графа знаний, в данной работе он рассматривается как совокупность нескольких онтологий, которые населены экземплярами. Когда граф знаний построен, для бизнес-акторов формируется набор определенных визуальных представлений. Самое распространенное представление – шахматка знаний (рис. 1) – отображает проекцию онтологии источников знаний на онтологию предметной области.

На пересечении этих проекций указаны значения параметров, которые также являются частью онтологии. Параметры разложены по категориям, связанным с жизненным циклом знаний. Более подробное описание каждого этапа разработки методологии, её апробации и анализа результатов картирования изложено ниже в соответствующих разделах.

Согласно [9] существуют принципиальные различия между организациями, которые придерживаются целевого (targeted) и комплексного (holistic) подхода к картированию знаний. Эти два подхода описаны далее.

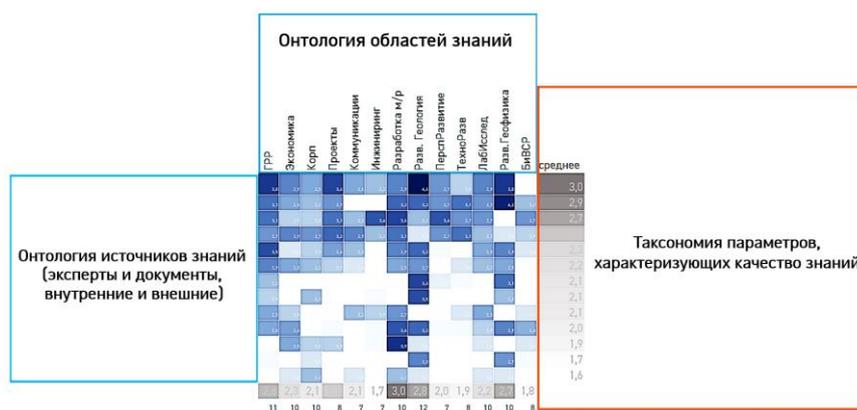


Рис. 1. Шахматка знаний как кросс-мэпинг нескольких онтологий

1. Разработка процесса картирования. Согласно исследованию [9] подходы, которые применяются для определения потребностей и возможностей организации в конкретной области знаний и/или в масштабах всего предприятия, можно разделить на два типа – целевой и комплексный. Ниже они кратко изложены в соответствии с указанным исследованием.

Целевой подход к картированию знаний заключается в фокусировке на определенных критически значимых областях, знания в которых обладают наибольшей полезностью и применимостью. Важно отметить, что целевой подход предполагает, что организация расставила приоритеты в своих потребностях в знаниях и определила риски до начала процесса картирования. Эта особенность делает целевое картирование знаний наиболее подходящим для организаций, которые уже имеют четкое представление о необходимости документирования конкретных экспертных знаний и опыта. Использование целевого подхода обеспечивает создание индивидуальных карт знаний по критически важным для организации областям. Такие карты далее совмещаются и связываются между собой для формирования общекорпоративного представления. Процесс картирования описан в Таблице 1.

Таблица 1. Пятиэтапный процесс картирования.
Составлено авторами на основе материалов [9]

Шаг	Результаты работ	Цель	Участники
1	Перечень ключевых областей знаний для картирования	Сфокусировать процесс картирования	Эксперты, команда управления знаниями (УЗ)
2	Матрица рисков утраты знаний	Приоритизировать области знаний по: 1) вероятности потери знаний 2) величине последствий	Руководители, специалисты HR, команда УЗ
3	Карта знаний	Формализовать знания с самым высоким риском	Команда УЗ, эксперты
4	Методология и план-график передачи знаний	Гарантировать, что знания эксперта будут сохранены и распространены	Команда УЗ, руководители,
5	Ежемесячные отчеты о ходе выполнения работ по передаче знаний		Исполнители мероприятий

Целевое картирование знаний предоставляет организациям возможность сфокусироваться на точном определении, фиксации и передаче накопленного опыта, однако оно не обеспечивает стратегических преимуществ, которые дает предприятиям и организациям комплексный подход. Комплексный подход направлен на отображение всех знаний, связанных с функциональным направлением, дисциплиной, бизнес-подразделением или предприятием.

Комплексный подход, предлагающий определенные стратегические преимущества. Цель такого подхода состоит в том, чтобы понять потребности и возможности в конкретной области знаний в масштабах всего предприятия. Комплексный подход обладает преимуществами, в сравнении с целевым картированием знаний. В частности, комплексные карты отображают:

- более широкое представление о рисках, пробелах и узких местах (bottlenecks) в знаниях;
- взаимозависимости и активы знаний, которые актуальны для множества групп;
- более широкий взгляд на то, какие подходы УЗ актуальны и необходимы организации, где и как их применять.

Важно отметить, что комплексное картирование знаний требует значительных временных затрат и ресурсов. Например, команде УЗ компании Transpetro понадобилось 18 месяцев, чтобы составить карту трех крупных направлений бизнеса. Выбор подхода зависит от ресурсов, от того, является ли приоритетом ширина или глубина исследования, а также от конкретных целей, которые необходимо достичь. Один из вариантов – начать создавать целевые карты, а затем перейти к более комплексному картированию знаний. Например, компания Goodyear запустила целевое картирование для экспертных знаний более десяти лет назад, но с тех пор добавила отдельную, более комплексную деятельность по картированию знаний, связанных с технологическими процессами. Эти карты помогают команде Goodyear по УЗ выявлять взаимозависимости и обеспечивать плавный переход проектов от одной части жизненного цикла продукта к другой.

Итоговый процесс картирования (рис. 2) был разработан как интеграция вышеописанных подходов, а также опираясь на работу (Balaid et. al 2016) и кейс Transpetro (APQC, 2018). Команда УЗ Transpetro создает карты для каждого макропроцесса в бизнес-области, используя пятиэтапный процесс для управления картированием знаний:

1. Определение знаний по областям.
2. Классификация критических знаний.
3. Определение людей с наличием критических знаний.
4. Приоритизация рисков.
5. Разработка плана по устранению рисков.

Коренным отличием от представленного в Таблице 1 процесса является то, что Transpetro выполняет анализ рисков на четвертом этапе. Благодаря комплексному подходу компания может подробно и всесторонне рассмотреть потенциальные риски, а не только те, о которых руководители уже знают.

В разработанном авторами процессе (рис. 2) уточнение областей знаний происходит итеративно.

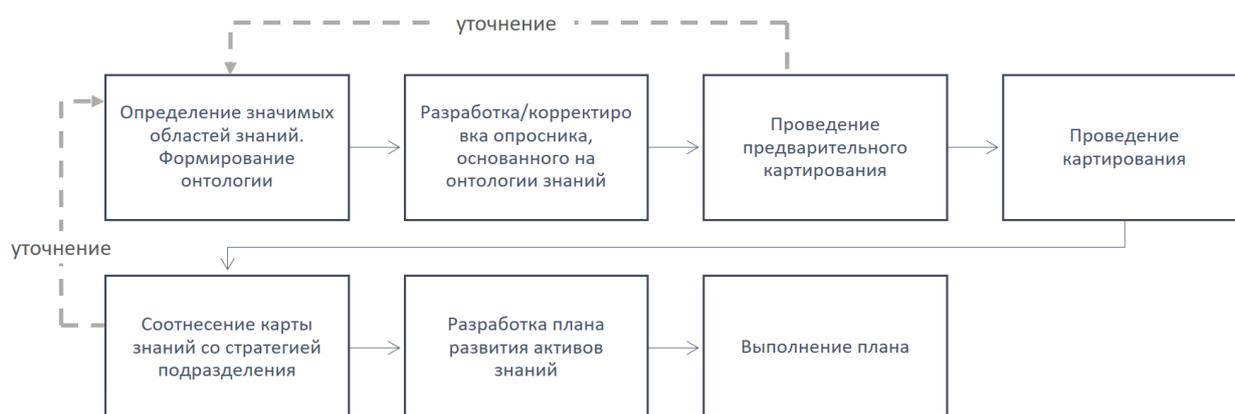


Рис. 2. Процесс картирования в наукоёмкой организации

2. Разработка онтологии картирования. В процессе разработки авторы столкнулись с вызовами, которые являются типовыми для такого рода работ в области инженерии знаний [10]. Основной проблемой при формировании системы интеллектуальной поддержки принятия решений на основе картирования знаний является разрозненность и неоднородность источников и областей знаний. Источники знаний, как таковые, и их типология не были ранее формализованы. Неоднородность проявляется в соседстве знаний различного уровня детализации, т.е. часть знаний описана только по функциям, а часть – описана и проработана до мельчайших деталей на пять уровней таксономии ниже. Выравнивание и расширение онтологии происходило совместно с рабочей группой экспертов.

Характерным является тот факт, что онтология не была формализована, а использовалась как набор связанных справочников и классификаторов. Это связано с необходимостью обеспечить человеко-читаемость на любом устройстве и в любой системе. При работе с экспертами предметной области наиболее эффективными оказались таблицы с функцией чек-листов. Для формирования опроса в информационной системе также необходимы были именно справочники и карты связей между ними, а не формальные owl-модели.

2.1. Кастомизированный классификатор знаний. Классификатор (с точки зрения инженерии знаний, это онтология) областей знаний разрабатывался итеративно. Первым шагом проводился пробный сбор данных, после чего категории расширялись и корректировались. Отдельное внимание было уделено классификатору знаний по ключевой проблематике, который разрабатывался с нуля. Поскольку, речь идет об узкоспециализированной деятельности, формального описания её составляющих ранее не существовало в компании. Первичный перечень был сформирован на основе частотного извлечения концептов из научных публикаций. Поскольку ключевая проблематика лежит в области активного научного поиска, удалось найти несколько сотен релевантных публикаций.

Выявленный перечень был обработан экспертами предметной области. Они выполнили следующие операции:

- Убрали синонимию и англицизмы.
- Обеспечили полноту: например, если есть понятия, касающиеся материка, то должны быть аналогичные, касающиеся шельфа.
- Выравнивали глубину полученного списка и привязали его к общему классификатору.

Далее полученный классификатор валидировался и дополнялся в ходе сбора данных – респонденты могли вносить свои значения и оставлять комментарии. Таким образом, была выявлена ранее не попадавшая в фокус внимания область – аномально высокое пластическое давление.

2.2. Оценка знаний сотрудника. На рисунке 3 представлен набор параметров и отображено формирование общего рейтинга. Он складывается из рейтинга по функции и общего рейтинга сотрудника. Функция «геологоразведочные работы» разделяется на несколько функциональных направлений, в рамках которых выделены отдельные области знаний. По каждой из областей знаний сотрудник отвечал бинарно: владею/не владею и указывал рекомендации – к кому он обратится за советом в выбранной области [11]. Рекомендации позволили выявить скрытых и подтвердить квалификацию существующих экспертов.



Рис. 3. Рейтинг знаний сотрудника как совокупность набора параметров

Вторая часть рейтинга – это набор параметров, определяющих базисный рейтинг сотрудника. Он включает основное и дополнительное образование, опыт (количество лет работы в отрасли), применение знаний в проектах, навыки обработки данных с помощью специализированного ПО, авторство документов.

Суммарно базисный рейтинг и рейтинг знаний по определенным областям формировали коэффициент, нормированный по шкале от 1 до 5. Это придаёт «шахматке» определенную глубину, отображающую не только сам факт того, что человек что-то знает, но и некую прокси-величину его экспертности. Это является некой альтернативой оценке компетенций, которая регулярно проводится, но требует больших затрат времени и не охватывает все возможные области знаний, фокусируясь на проверке соответствия сотрудника исключительно текущей должности.

3. Описание выборки и сбор данных. В исследуемом подразделении компании неоднократно проводилось картирование, охватывавшее несколько сотен человек. При этом было установлено, что такое количество респондентов распадается на 3-5 кластеров, каждый из которых сфокусирован на отдельной предметной области. Детальная проработка при этом должна проводиться по каждому отдельно взятому кластеру.

В описываемом случае картирование было сфокусировано на группе сотрудников, работающих над одной общей проблематикой – трудноизвлекаемыми запасами выбранного региона. В опросе участвовал 41 сотрудник картируемого подразделения, таким образом, доля ответивших составила 95,3%, из них 27 сотрудников относятся к функциональной дирекции (93% доля ответивших) и 14 сотрудников – из научного центра (100% доля ответивших).

Предметом картирования стал следующий набор областей знаний:

1. Бурение и внутрискважинные работы.
2. Геология и разработка.

3. Геологоразведочные работы.
4. Инжиниринг. Реинжиниринг.
5. Перспективное развитие.
6. Проектные сервисы.
7. Общекорпоративные области, такие, как промышленная безопасность, управление изменениями и постоянное совершенствование, закупки и снабжение, управление поставщиками и подрядчиками, общая цифровая грамотность, управление командами, договорная деятельность.

Ввиду многодисциплинарного характера изучаемой проблематики, при картировании были учтены различные направления деятельности сотрудника при выполнении работ:

- Функциональные знания.
- Углубленные профильные знания.
- Владение специализированным ПО.
- Обладание знаниями о контрагентах/партнерах.
- Знания, полученные в результате работы над проектами.

Программное обеспечение, направление деятельности партнеров и тематика проектов были соотнесены с основными категориями знаний

Сбор данных для цифровой карты осуществлялся с помощью опроса сотрудников и дополнялся информацией, предоставленной подразделением. Опрос разрабатывается на основе апробированных научных работ [11,12] с учетом специфики Компании. В апробации методики участвовали 98 сотрудников в первую волну и 41 – во вторую. В начале каждой волны до 10 человек участвовали в тестировании опросника, после чего в него вносились корректировки.

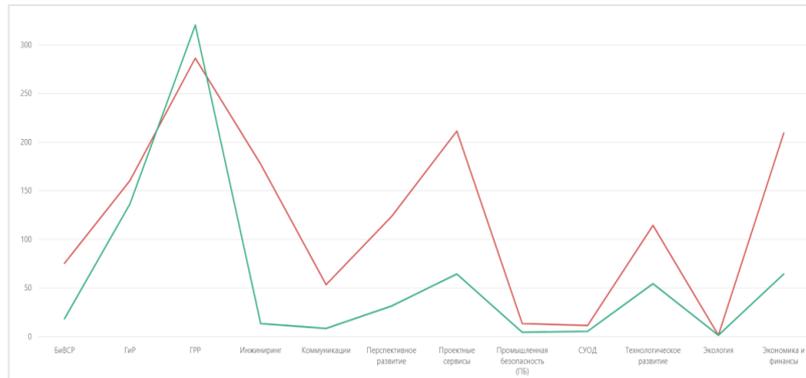
За счет таксономической организации областей знаний, при которой сотрудник мог выбрать только значимые для него категории и только по ним отвечать на детализирующие вопросы, стало возможным охватить картированием 254 области знаний. Также этому способствовал вышеупомянутый бинарный выбор – авторы отказались от идеи самооценки сотрудника по пятибалльной шкале в пользу множественного выбора из предоставленного перечня. Для сравнения, типовая оценка компетенций для определения соответствия должности включает не более 30 областей и занимает несколько часов.

4. Результаты. Результатом картирования стало построения ряда визуальных представлений, основным из которых является вышеописанная «шахматка» знаний, где в строках указаны области знаний – это нижний уровень разработанного классификатора, а в столбцах представлены сотрудники, источники знаний. На рисунке 3 представлен фрагмент карты знаний. Всего в ней 254 строки и 39 столбцов (2 респондента заполнили только ту часть опросника, которая посвящена узкоспециализированным знаниям – см. ниже).

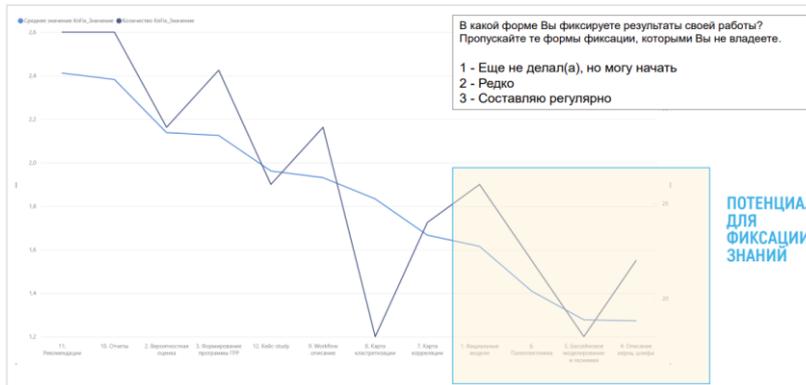
Отдельное внимание уделено тем областям, которые сейчас не входят в круг обязанностей сотрудника, но по которым у него есть опыт. Респондент мог выбрать ответ «знаниями владею, но сейчас не выполняю задачи по этому направлению; готов(а) начать работу по нему». Этот ответ позволил выявить скрытый потенциал диверсификации решаемых задач подразделения.

Сравнение проектных команд по объему знаниевых активов и системности фиксации знаний в документах, представленное на рисунке 4, позволяет проводить инвестиционную оценку и выявлять потенциал роста команд. На рисунке 5 представлен фрагмент карты узкоспециализированных знаний предметной области. Полная карта состоит из 23 строк и 41 столбца. Узкоспециализированные знания, относящиеся к ключевой проблематике предметной области, измерялись по шестибалльной шкале от «0» означающего отсутствие знаний в этой области до «5», соответствующего высокому уровню экспертизы. Сделав шаг назад и

взглянув на карту знаний организации в целом, команда УЗ может выявить области, в которых знания без необходимости дублируются или где они не передаются другим группам, которые могли бы извлечь из этого пользу.



а)



б)

Рис. 4. Сравнение двух проектных команд по параметрам: а) количество областей знаний по различным направлениям, которые сотрудники указали как свои ключевые, б) потенциал фиксации знаний в документах

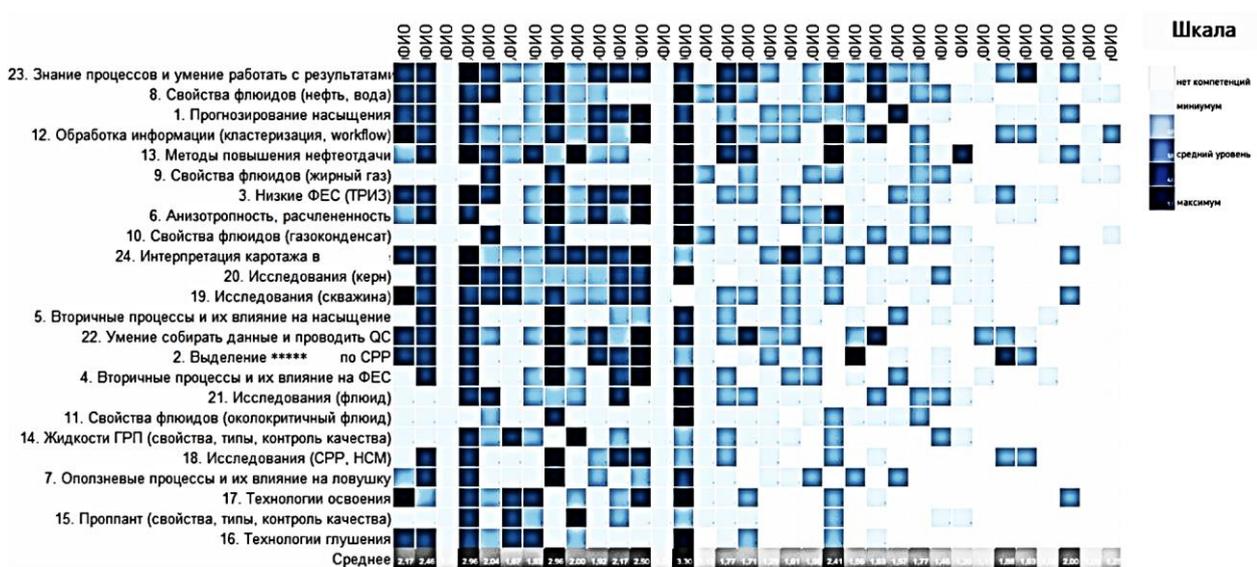


Рис. 5. Карта узкоспециализированных знаний предметной области

Потенциал для принятия управленческих решений на основе представленной карты включает в себя: выявление сотрудников, нуждающихся в консультационной поддержке и раз-

витии определенных навыков; активация высококвалифицированных экспертов; формирование плана личной передачи знаний (коучинг, шэдоуинг¹, менторство). На рис. 6 представлен дашборд (информационная панель) знаний отдельного сотрудника.

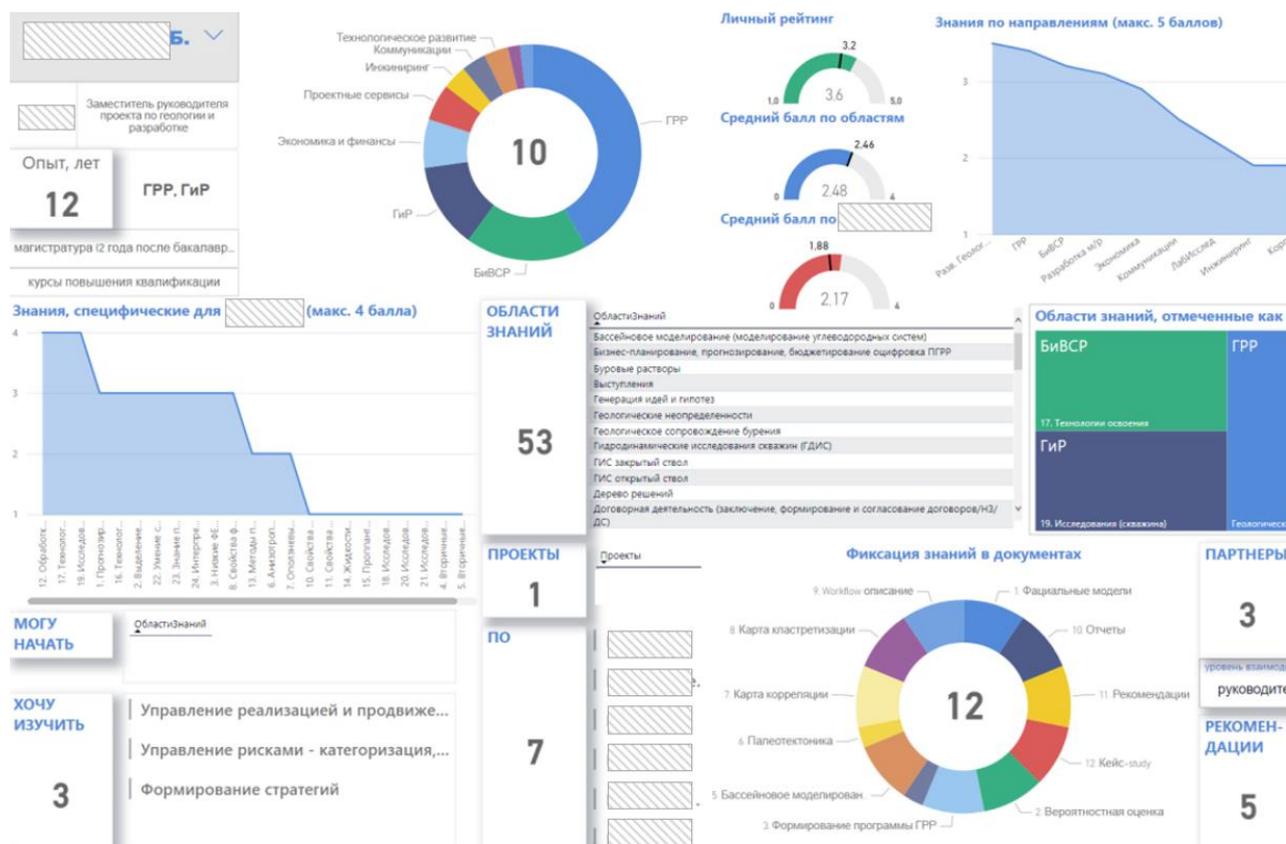


Рис. 6. Дашборд знаний отдельного сотрудника

По итогам картирования выявлены наиболее активные сотрудники, стремящиеся развиваться и лидировать в определенных направлениях. Представлена расширенная модель знаний, охватывающая всех ключевых сотрудников картируемого подразделения и ключевые направления деятельности.

Итогом внедрения методики стал набор рекомендаций по развитию интеллектуального капитала картируемого подразделения. В частности, произведен анализ проектов с точки зрения их ценности как источника уникального опыта (см. рис. 5).

Созданы переиспользуемые системы организации знаний. Справочники, перечни, тезаурусы, которые были разработаны в ходе картирования, поэтапно интегрируются в корпоративную систему управления знаниями. Они включаются в информационные модули, такие, как извлеченные уроки, вопросы эксперту и др.

Как подготовительный этап перехода к системе интеллектуального принятия решений на основе данных картирования рекомендованы к разработке:

1. План системной передачи знаний сотрудников.
2. План системной фиксации знаний в документах.
3. План формирования авторских разработок и других интеллектуальных активов.
4. Базовая система для работы со знаниями о партнерах.

¹ Шедоуинг (от англ. shadow, тень) – «теневой метод». Метод обучения на рабочем месте, предполагающий наблюдение за работой опытного сотрудника с целью совершения аналогичных действий в своей работе. Обучаемый не комментирует происходящего, не задает уточняющих вопросов, а наблюдает за действиями опытного сотрудника.

Сформирован план развития интеллектуального потенциала подразделения в части знаний о проектах (рис. 7).



Рис. 7. Формирование плана развития интеллектуального потенциала подразделения в части знаний о проектах

Заключение. Систематическое картирование, категоризация и сравнительный анализ знаний организации позволяют принимать информированные управленческие решения. Организация может получить большую отдачу от усилий по развитию, сгруппировав познавательный опыт в «критическую научающую массу» вокруг некоторых стратегических областей знания [1]. Более того, когда сотрудник переходит из одного подразделения в другое и/или меняет направление деятельности, то значимый блок его знаний не уходит из поля зрения коллег и руководителей.

Карты знаний служат руководством к тому, какие знания должны быть получены, от кого и для кого. Важным шагом любого подхода является то, что организация выполняет оценку рисков или другое мероприятие по установлению приоритетов, чтобы определить, какие области требуют более подробного картирования, документирования или передачи.

Результатом проведения картирования является формирование общего смыслового поля, обеспечивающего взаимопонимание и согласованный обмен знаниями внутри профессионального сообщества.

Благодарности. Работы Гринберг Э.Я. по данной статье выполнены на средства гранта РФФИ №23-21-00168 «Методология и технология разработки цифровых карт знаний для учебных и научных коллективов (МЕТАКАРТА)», 2023-2024.

Список источников

1. Zack M.H. Knowledge and strategy. Routledge, 2009, pp. 17-18.
2. Митякова О.И. Пути совершенствования производственной системы нефтегазовой компании / О.И. Митякова, А.В. Цеханский // Инновации и инвестиции, 2023. – №4. – С. 82-85.
3. Наплёкова Т.К. Практические аспекты повышения эффективности производственной стратегии промышленных предприятий региона / Т.К. Наплёкова, Т.Н. Пигина // Вестник СибГИУ, 2023. – № 1 (43). – С. 104-115.
4. Яковлева Е.А. Модель анализа материальных и нематериальных факторов эффективности в системе стратегического планирования на основе совместного причинно-следственного картирования и фреймового представления знаний / Е.А. Яковлева, М.М. Гаджиев М.М., Шарич Э.Э., Яковлева Д.Д. // Экономика, предпринимательство и право, 2021. – №1. – С. 145-155.
5. Антюхов В.И. Моделирование процесса интеллектуальной поддержки деятельности должностных лиц центров управления в кризисных ситуациях МЧС России при принятии управленческих решений / В.И. Антюхов, Н.В. Остудин // Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России, 2017. – № 2. – С. 78-93.

6. Заковоротный В.Л. Синергетический подход к оценке человеческого капитала предприятия / В.Л. Заковоротный., М.Б. Флек, Е.А. Угнич // Научно-технические ведомости СПбГПУ. Экономические науки, 2019. – Т. 12. – № 6. – С. 161-173.
7. Kudryavtsev D.V., Gavrilova T.A., Grinberg E.Y., Kubelskiy M.V. Map of the Maps: Conceptualization of the knowledge maps. BIR Workshops 2022: 14-23, Rostock, Germany, available at: <https://ceur-ws.org/Vol-3223/paper2.pdf>.
8. Карлик А.Е. Совместное когнитивное картирование -метод обеспечения междисциплинарных инновационных проектов меганауки / А.Е. Карлик, В.В. Платонов, С.А. Кречко // Экономическая наука современной России, 2018. – № 4 (83). – С. 65-84.
9. Harper M., Trees L. Knowledge mapping in action. How organizations map their knowledge and learn from the results. An APQC best practices report, 2018, available at: <https://www.apqc.org/resource-library/resource-listing/knowledge-mapping-action-best-practices-report> (accessed: 08/20/2023).
10. Гаврилова Т.А. Визуально-аналитическое мышление и интеллект-карты в онтологическом инжиниринге / Т.А. Гаврилова, Э.В. Страхович // Онтология проектирования, 2020. – Т. 10. – № 1 (35). – С. 87-99.
11. Liebowitz J. Linking social network analysis with the analytic hierarchy process for knowledge mapping in organizations. Journal of knowledge management, 2005, vol. 9, no. 1, pp. 76-86.
12. Balaid A. [et al.] Knowledge maps: A systematic literature review and directions for future research. International journal of information management, 2016, vol. 36, no. 3, pp. 451-475.

Гринберг Эльвира Яковлевна. Главный специалист лаборатории цифрового моделирования СПбГУ. SPIN: 7438-2164, AuthorID: 876459, ORCID 0000-0001-8871-1349, st057454@student.spbu.ru, 199034, Россия, Санкт-Петербург, Университетская наб., д. 7–9

Верзин Евгений Андреевич. Руководитель направления по развитию продукта центра по управлению знаниями, ООО «Газпромнефть НТЦ». SPIN: 7772-6134, AuthorID: 1210569, ORCID 0009-0002-0751-0073, Verzin.EA@gazpromneft-ntc.ru, 190000, Россия, г. Санкт-Петербург, Набережная реки Мойки, д. 75–79, литер Д.

UDC 005.94

DOI: 10.25729/ESI.2023.32.4.013

Development and testing of the knowledge mapping methodology of a knowledge-intensive organization

Elvira Ya. Grinberg¹, Evgeny A. Verzin²

¹ St. Petersburg State University, Russia, St. Petersburg, st057454@student.spbu.ru

² LLC “Gazpromneft Science & Technology Centre”, Russia, St. Petersburg

Abstract. Systematic mapping, categorization and comparative analysis of the organizational knowledge allows executives to make informed managerial decisions. The fragmentation and heterogeneity of knowledge sources, as well as the lack of unification of their assessment is key problem in forming a system of intellectual decision support based on knowledge mapping. This paper presents the development and testing of a knowledge mapping methodology as a response to the described problem. A distinctive feature of the presented approach is the methodological reliance on ontological engineering. The ontology was developed and proposed for updating to the working group. It took into account the following areas: functional knowledge, in-depth profile knowledge, possession of specialized software, knowledge of contractors, knowledge gained as a result of project work. Data collection for the digital map was carried out using a survey of employees and supplemented with information provided by the sub-division.

Keywords: Knowledge map, ontology, competence mapping, knowledge dashboard

Acknowledgements: The works of E. Grinberg on this article were carried out at the grant funds of Russian Science Foundation (project number 23-21-00168) “Methodology and technology for developing digital knowledge maps for education and research teams (METAКARTA)”, 2023-2024.

References

1. Zack M.H. Knowledge and strategy. Routledge, 2009, pp. 17-18.
2. Mityakova O.I., Cekhanskij A.V. Puti sovershenstvovaniya proizvodstvennoj sistemy neftegazovoj kompanii [Ways to improve the production system oil and gas company]. Innovacii i investicii [Innovation and investment], 2023, no. 4, pp. 82-85.
3. Naplyokova T.K., Pigina T.N. Prakticheskie aspekty povysheniya effektivnosti proizvodstvennoj strategii promyshlennyh predpriyatij regiona [Practical aspects of increasing the efficiency of the production strategy of industrial enterprises region]. Vestnik SibGIU [Bulletin of the Siberian State Industrial University], 2023, no. 1 (43), pp. 104-115.
4. Yakovleva E.A., Gadzhiev M.M., Shaich E.E., Yakovleva D.D. Model' analiza material'nyh i nematerial'nyh faktorov effektivnosti v sisteme strategicheskogo planirovaniya na osnove sovmestnogo prichinno-sledstvennogo kartirovaniya i frejmovogo predstavleniya znaniy [Model of analysis of material and intangible efficiency factors in the strategic planning system based on joint cause-and-effect mapping and frame representation of knowledge]. Ekonomika, predprinimatelstvo i pravo [Journal of economics, entrepreneurship and law], 2021, no. 1, pp. 145-155.
5. Antyuhov V.I., Ostudin N.V. Modelirovanie processa intellektual'noj podderzhki deyatel'nosti dolzhnostnyh lic centrov upravleniya v krizisnyh situacijah MCHS Rossii pri prinyatii upravlencheskih reshenij [Modeling of the process of intellectual support of activities of officials of control centres in crisis situations of emercom of Russia in making management decisions]. Vestnik Sankt-Peterburgskogo universiteta Gosudarstvennoj protivopozharnoj sluzhby MCHS Rossii [Bulletin of St. Petersburg University of the State Fire Service of the Ministry of Emergency Situations of Russia], 2017, no. 2, pp. 78-93.
6. Zakovorotnyj V.L., Flek M.B., Ugnich E.A. Sinergeticheskij podhod k ocenke chelovecheskogo kapitala predpriyatiya [Synergetic approach to assessment of human capital in enterprises]. Nauchno-tehnicheskie vedomosti SPbGPU. Ekonomicheskie nauki [St. Petersburg State Polytechnical university journal. Economics], 2019, vol. 12, no. 6, pp. 161-173.
7. Kudryavtsev D.V., Gavriloza T.A., Grinberg E.Y., Kubelskiy M.V. Map of the Maps: Conceptualization of the Knowledge Maps. BIR Workshops 2022: 14-23, Rostock, Germany, available at: <https://ceur-ws.org/Vol-3223/paper2.pdf>.
8. Karlik A.E., Platonov V.V., Krechko S.A. Sovmestnoe kognitivnoe kartirovanie -metod obespecheniya mezhdisciplinarnykh innovacionnykh proektov meganauki [Participative cognitive mapping -a method to support the interdisciplinary innovative projects of megascience]. Ekonomicheskaya nauka sovremennoj Rossii [Economics of contemporary Russia]. 2018, no. 4 (83), pp. 65-84.
9. Harper M., Trees L. Knowledge mapping in action. How organizations map their knowledge and learn from the results. An APQC best practices report, 2018, available at: <https://www.apqc.org/resource-library/resource-listing/knowledge-mapping-action-best-practices-report> (accessed: 08/20/2023).
10. Gavriloza T.A., Strahovich E.V. Vizual'no-analiticheskoe myshlenie i intellekt-karty v ontologicheskom inzhiniringe [Visual analytical thinking and mind maps for ontology engineering]. Ontologiya proektirovaniya [Ontology of designing], 2020, vol. 10, no. 1 (35), pp. 87-99.
11. Liebowitz J. Linking social network analysis with the analytic hierarchy process for knowledge mapping in organizations. Journal of knowledge management, 2005, vol. 9, no. 1, pp. 76-86.
12. Balaid A. et al. Knowledge maps: A systematic literature review and directions for future research. International journal of information management, 2016, vol. 36, no. 3, pp. 451-475.

Grinberg Elvira Yakovlevna. Lead specialist of the Digital modeling laboratory of St. Petersburg University. SPIN: 7438-2164, AuthorID: 876459, ORCID: 0000-0001-8871-1349, st057454@student.spbu.ru, 199034, Russia, Saint Petersburg, Universitetskaya nab., 7-9.

Verzin Evgeny Andreevich. Head of product development in knowledge management department, LLC "Gazpromneft Science & Technology Centre". SPIN-code: 7772-6134, AuthorID: 1210569 ORCID: 0009-0002-0751-0073, Verzin.EA@gazpromneft-ntc.ru. 190000, Russia, Saint Petersburg, Moika river emb, 75-79 letter D.

Статья поступила в редакцию 28.08.2023; одобрена после рецензирования 29.11.2023; принята к публикации 11.12.2023.

The article was submitted 08/28/2023; approved after reviewing 11/29/2023; accepted for publication 12/11/2023.

УДК 004.94

DOI: 10.25729/ESI.2023.32.4.014

Мониторинг и накопление данных в информационно-прогностической системе ГеоГИПСАР

Гасан Виктор Святославович, Абасов Николай Викторович, Осипчук Евгений Николаевич

Институт систем энергетики им. Л.А. Мелентьева СО РАН, Россия, Иркутск,
viktor-gasan@yandex.ru

Аннотация. В статье рассматриваются исходные данные, их структура и методы обработки, применяемые в информационно-прогностической системе ГеоГИПСАР, которая была разработана в Институте систем энергетики им. Л.А. Мелентьева (ИСЭМ) СО РАН и продолжает развиваться. Она успешно использует данные современных глобальных климатических моделей (таких как NOAA, GPCP) для прогнозирования притока воды в крупные водохранилища и озера с оценкой их уровневых режимов. В статье приводится описание структур и типов показателей глобальных климатических моделей, используемых на практике при прогнозировании стока в бассейнах рек. Рассмотрены основные шаги для сбора данных, их обработки и анализа. Демонстрируются результаты работы системы с использованием разработанного визуализатора карт. Отмечается необходимость проведения исследований по прогнозированию полезного притока для эффективного управления уровнем режимом озера Байкал. По результатам работы планируется расширение системы с включением веб-интерфейса для анализа и формирования прогностических показателей.

Ключевые слова: глобальные климатические модели, реанализ, мониторинг, информационно-прогностическая система

Цитирование: Гасан В.С. Мониторинг и накопление данных в информационно-прогностической системе ГеоГИПСАР / В.С. Гасан, Н.В. Абасов, Е.Н. Осипчук // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 160-169. – DOI:10.25729/ESI.2023.32.4.014.

Введение. За последние десятилетия наблюдаются значительные успехи в мониторинге и прогнозировании климата. Однако надежные универсальные методики их применения на практике отсутствуют. В условиях изменения климата важной задачей становится планирование режимов крупных водохранилищ и регулируемых озер, например, оз. Байкал. При долгосрочном планировании режимов обычно используются оценки по норме стока на основе накопленных статистических данных. В связи с ужесточением современных экологических требований, как на уровневые режимы оз. Байкал, так и на расходы воды через Иркутскую ГЭС, а также появлением экстремальных и продолжительных маловодных и многоводных периодов, проведение оценок будущих режимов с использованием только статистических данных становится ограниченным, что требует применения новых подходов.

Для обработки и анализа геоклиматических данных в ИСЭМ СО РАН разработана и развивается информационно-прогностическая система ГеоГИПСАР [1–6], которая использует универсальную базу с файлами многомерных сеточных данных специализированных форматов с включением в них различных агрегированных показателей. Система основана на компонентном подходе с инструментальными средствами развития и добавления новых компонентов.

Сложность мониторинга и обработки глобальных климатических данных связана как с большими объемами исходных данных (до десятков Гб в день), включающих множество атрибутов и показателей с разными временными интервалами по всему земному шару, так и необходимостью учета наличия пропусков и ошибок в данных, верификации и согласования данных из различных источников. Кроме того, как правило, данные из различных центров используют разные форматы, которые имеют сложную структуру и описание, что приводит к проблеме их оперативной обработки.

1. Исходные данные и их структура. Предлагаемый подход к мониторингу фактических показателей параметров атмосферы в бассейне оз. Байкал с помощью ГеоГИПСАР основан на ретроспективном анализе данных с центра реанализа.

Центр данных реанализа включает:

- ERA5/NCMWF [7] – данные реанализа атмосферы и глобального климата NCMWF пятого поколения, охватывающий период с января 1940 г. по настоящее время;
- ECMWF ERA-40 [7] – данные реанализа ECMWF глобальной атмосферы и состояния поверхности за 45 лет (с сентября 1957 г. по август 2002 г.);
- NOAA/NCEP [8] – данные реанализа Национального управления океанических и атмосферных исследований (США) с координатной сеткой $2,5 \times 2,5^\circ$;
- GPCC [9-11] – мировой центр данных по осадкам (Германия) с месячным разрешением за период 1900–2022 гг. и координатной сеткой 1×1 .

Центр точечных данных включает:

- Weather and Climate [12] – оперативные метеоданные по пунктам;
- Aisori-M [13] – архивы метеоданных по пунктам;
- Predictive ensembles [14] – прогностические ансамбли;
- AIS GMVO [15] – архивы данных по водным объектам;
- Solar activity [16] – данные по солнечной активности.

В системе ГеоГИПСАР, с использованием информации из мировых центров данных и глобальных прогностических моделей, постоянно отслеживаются глобальные и региональные климатические изменения с формированием различных вероятностных прогностических оценок по притокам для исследуемых бассейнов. Реализован набор стандартных конверторов и шаблонов обработки различных выборок данных из открытых источников. Ежедневно производится опрос сайтов на обновление данных с загрузкой необходимых файлов, преобразованием к нужному формату и пересылкой в соответствующее хранилище. В ГеоГИПСАР используются как точечные, так и пространственно-распределенные данные.

Точечные данные классифицируются:

- по пунктам метеорологических наблюдений;
- по типам данных (температура, осадки, давление, др.);
- по времени (годовые, месячные, декады, суточные, срочные).

Пространственно-распределенные данные включают:

- текстовые форматы;
- формат NetCDF [17];
- формат HDF5 [18];
- формат GRIB [19].

NetCDF – формат файлов, ориентированный на многомерные массивы данных, широко используемый в ГИС. Он был принят и согласован Открытым геопространственным консорциумом (OGC) в качестве основного стандарта кодирования. Почти любой тип пространственных данных может быть сохранен в файле NetCDF. Формат часто используется в данных об атмосфере и океане. Одним из основных преимуществ формата является его самоописание – в самом файле содержатся метаданные, описывающие его структуру.

HDF – формат файлов, широко используемый в сообществе дистанционного зондирования. Как и в NetCDF, большая часть данных внутри файла HDF хранится в виде многомерных массивов. Одно из различий между этими форматами заключается в том, что

HDF поддерживает иерархическую структуру внутри файла. Данные в файле могут быть организованы в группы. Группы могут содержать один или несколько наборов данных. Каждый из этих наборов данных может содержать многомерные данные.

Существует три разных типа файлов HDF: HDF4, HDF5 и HDF-EOS. Устаревший формат HDF4 не поддерживал файлы размером более 2 Гб и имел ограничения в структурах ввода и вывода. HDF5 не имел этих ограничений, а также включал множество улучшений. HDF-EOS – специальная реализация HDF, оптимизированная для хранения больших объемов данных спутникового дистанционного зондирования.

Общая модель данных CDM (Common Data Model) [17] (рис. 1) используется в специализированных форматах хранения данных NetCDF и HDF5. Основное её назначение – представление многомерных численных массивов (рис. 1). Файл имеет определенную группу данных верхнего уровня. Каждая группа может содержать одну или несколько именованных подгрупп, определяемые пользователем типы, переменные, измерения и атрибуты. Переменные также имеют атрибуты и могут иметь общие измерения, указывающие на общую сетку. Один или несколько размеров могут иметь неограниченную длину. Переменные и атрибуты имеют 1 из 12-ти примитивных типов данных или 1 из 4-х пользовательских типов.

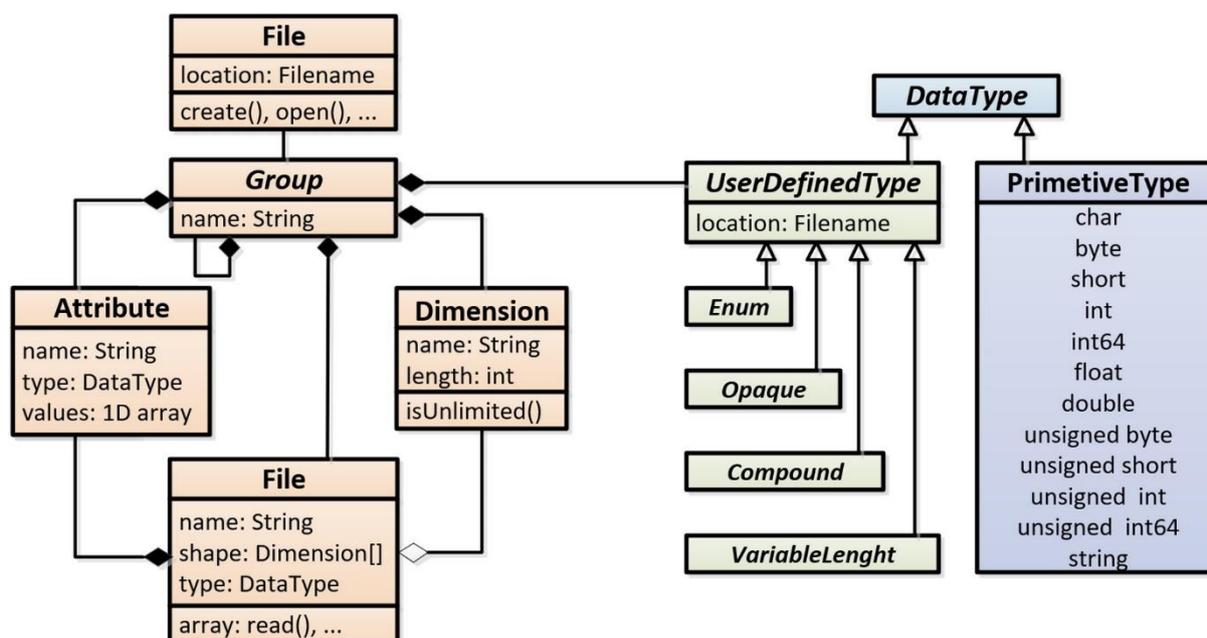


Рис. 1. Модель данных CDM

GRIB (англ. GRId in Binary – сетка в двоичном виде) – является форматом пространственных данных с применением математических методов сжатия. Формат широко используется для метеорологических данных и поддерживается Европейским центром среднесрочных прогнозов погоды (ECMWF) [7], Национальным управлением океанических и атмосферных исследований (NOAA) [8], Национальным центром атмосферных исследований (NCAR) [20].

В настоящее время используются 2 версии формата: GRIB1 с 1994 г. и GRIB2 с 2003 г. GRIB1 применяется для управления двумерными атмосферными данными с географической привязкой на поверхности сетки. GRIB2 расширяет возможности GRIB1 для 4D пространственно-временных данных и прогностических ансамблей. GRIB2 расширяет возможности описания метеорологических параметров, любых связанных статистических обработок и их метаданных, также включает возможности сжатия данных.

Файл формата GRIB2 состоит из нескольких секций [21] (рис. 2):

1. Информации о метаданных: содержит информацию о типе файла, способе кодирования и дополнительную мета-информацию.
2. Идентификации: содержит описание формата файла и данных, включая номер версии GRIB и их размер.
3. Определения структуры: включает содержание и структуру данных, например, тип данных, интервал времени, переменные и т.д.
4. Данных: содержит сами данные в двоичном формате, может содержать как сжатые, так и несжатые данные.
5. Карты данных (битовый массив): содержит информацию о наличии или отсутствии данных для каждого элемента в массиве, что позволяет сократить их объем.



Рис. 2. Секции данных в формате GRIB2

2. Используемые методы и технологии. Схема преобразования данных из различных центров в системе ГеоГИПСАР изображена на рис. 3.

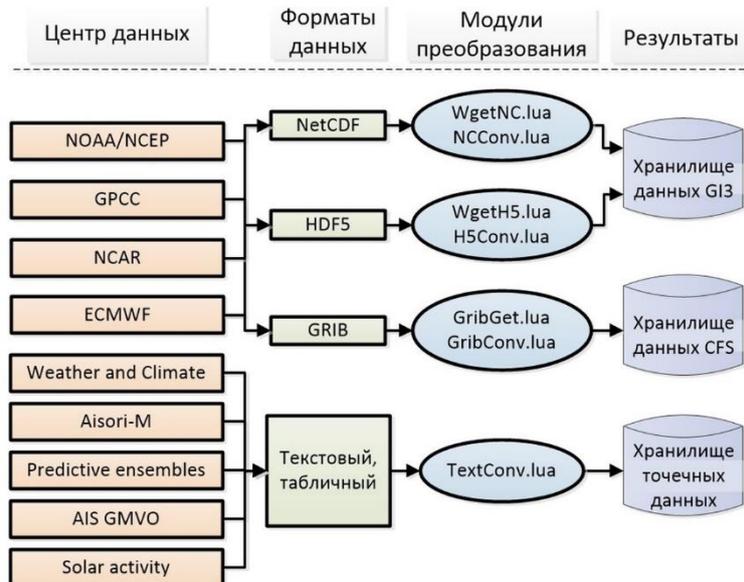


Рис. 3. Преобразование данных из различных центров в системе ГеоГИПСАР

Данные с пунктов наблюдения в бассейнах водосбора рек позволяют проверить и уточнить данные ретроспективного анализа. Для эффективности дальнейшего анализа исходные данные из форматов NetCDF и HDF5 конвертируются в специализированный GIZ-формат, что позволяет хранить в одном файле не только пространственные данные за весь

период наблюдения, но и включать в себя показатели среднего, минимального и максимальные значения для каждой даты в этом файле. Данные используются для суточного и месячного разрешения (рис. 4).

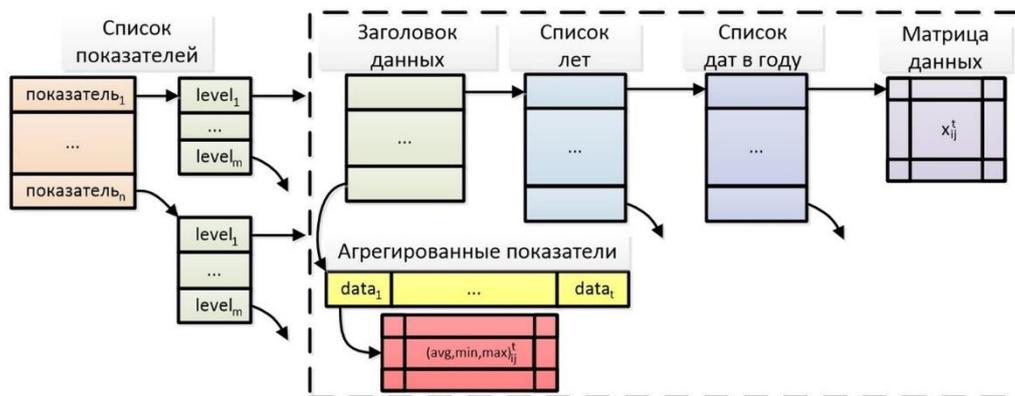


Рис. 4. Структура геоклиматических фактических показателей G_{i3}

Разработана технология преобразования ансамблей прогнозов (в формате GRIB) в специализированные файлы формата CFS, формируемых глобальной климатической системой CFSv2 [22–25] (рис. 5). Она включает среднесуточные прогностические показатели с прямым доступом к их элементам, наполнение базы прогностических ансамблей, а также технологию для обработки этих файлов и получение прогностических оценок распределения температуры, интенсивности осадков и приземного давления.

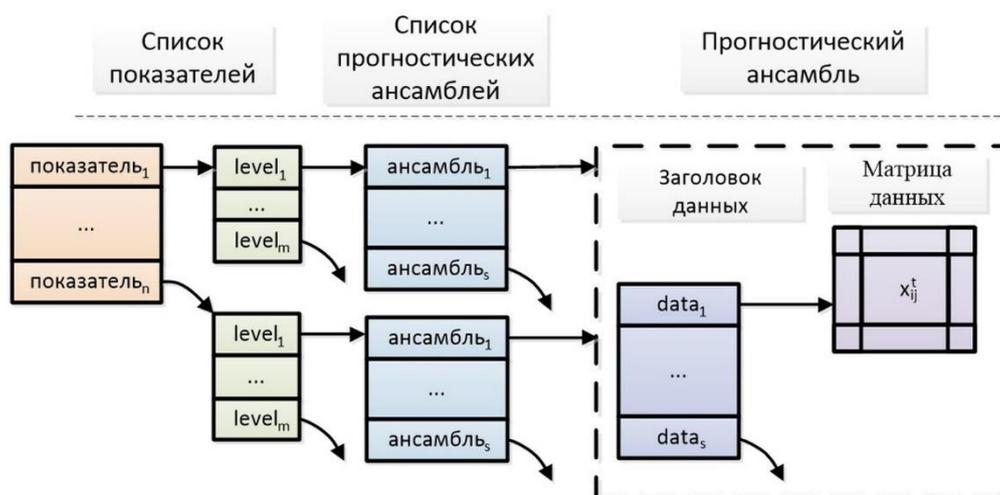


Рис. 5. Структура прогностических ансамблей CFSv2 в системе ГеоГИПСАР

CFSv2 (Climate Forecast System версии 2) – прогностическая система, разработанная в Центре моделирования окружающей среды (ЕМС), который входит в организацию NCEP. Система учитывает динамику изменения атмосферной циркуляции, океанских течений и льдов Арктики и Антарктики. На её основе ежедневно уточняются прогностические ансамбли состояния атмосферы и океана с временным интервалом от нескольких часов до 9 месяцев для всего земного шара.

Используемый в модели ансамблевый подход позволяет формировать вероятностные оценки состояния атмосферы на длительную перспективу. Модель является полностью связанной и представляющей взаимодействие между атмосферой Земли, океанами, сушей и морским льдом. Данная модель функционирует в NCEP с марта 2011 г. и предоставляет в открытом доступе прогностические данные по всему земному шару.

Для возможности оперативной работы с различными источниками и обновления единой базы данных в ГеоГИПСАР используются следующие шаги (рис. 6):

1. С помощью управляющей системы, используя свободные консольные утилиты для загрузки (CURL, WGET), автоматически скачиваются последние данные из различных центров (NOAA/NCEP, GPCP, NCAR, ECMWF);
2. Под управлением компонента LuaISEM системы ГеоГИПСАР, исходные файлы декодируются с использованием стандартных утилит NCDUMP, H5DUMP, WGRIB для выбора необходимых данных;
3. Полученные данные конвертируются в форматы базы ГеоГИПСАР (GI3, CFS) и записываются в хранилище.

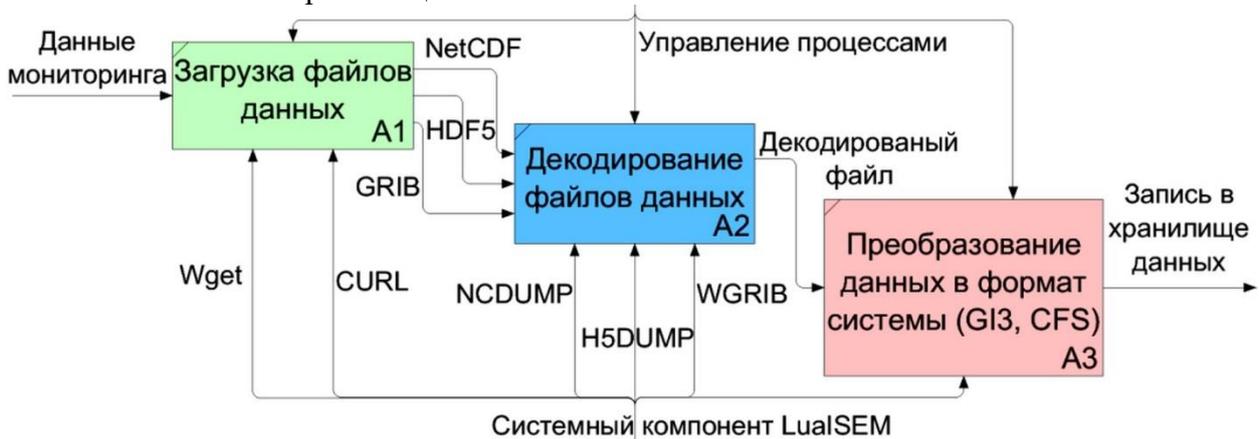


Рис. 6. Функциональная схема мониторинга и накопления данных в системе ГеоГИПСАР

Поддержка целостности данных осуществляется через мониторинг и обновление их состояния в центрах, сравнивая с состоянием баз данных в системе ГеоГИПСАР.

3. Использование ГеоГИПСАР для моделирования уровенных режимов оз. Байкал. При наполнении и сработке воды в оз. Байкал важным является исследование естественных режимов регулирования озера (до строительства Иркутской ГЭС), чтобы определить возможность максимального соответствия зарегулированных и природных (естественных) режимов. В современных условиях невозможно нарушить существующие требования по техническим параметрам работы Иркутской ГЭС, водохозяйственным, социально-экономическим и другим ограничениям в верхнем и нижнем бьефах. Даже при минимальном выполнении требований к зарегулированным режимам необходимо учитывать зависимости изменения уровня озера от расхода ГЭС, ограничения минимального расхода 1300 м³/с по требованию водозаборов, максимально допустимые расходы 4000–4500 м³/с по условиям незатопления прибрежных территорий в нижнем бьефе.

С помощью распределений прогностических показателей (осадки, температура, давление, геопотенциал, скорость и циркуляция атмосферы) в бассейнах водосбора оз. Байкал уточняется полезный приток, на основе которого моделируется динамика изменения его уровенного режима. Выделение на климатических картах границ бассейнов водосбора рек озера позволяет определить годы-аналоги по заданной мере близости в виде суммы среднеквадратичных отклонений с весовыми коэффициентами по выделенной области. На основе их формируются наиболее вероятные прогностические показатели полезного притока.

Высокие и низкие показатели полезного притока в оз. Байкал коррелируют с водностью р. Селенга. В случаях совпадения высокой или низкой водности с северными (Верхняя Ангара) и восточными реками (Баргузин) создаются условия появления экстремальной водности на озере.

Разработанная система позволяет формировать пространственные карты по нескольким ансамблям моделей за указанный период и по заданным климатическим показателям, например, геопотенциалу (рис. 7а); температуре (рис. 7б); давлению (рис. 7в); осадкам (рис. 7г). Распределение скоростей циркуляции атмосферы и геопотенциала на разных изобарических поверхностях позволяет визуализировать циклоническую и антициклоническую активность.

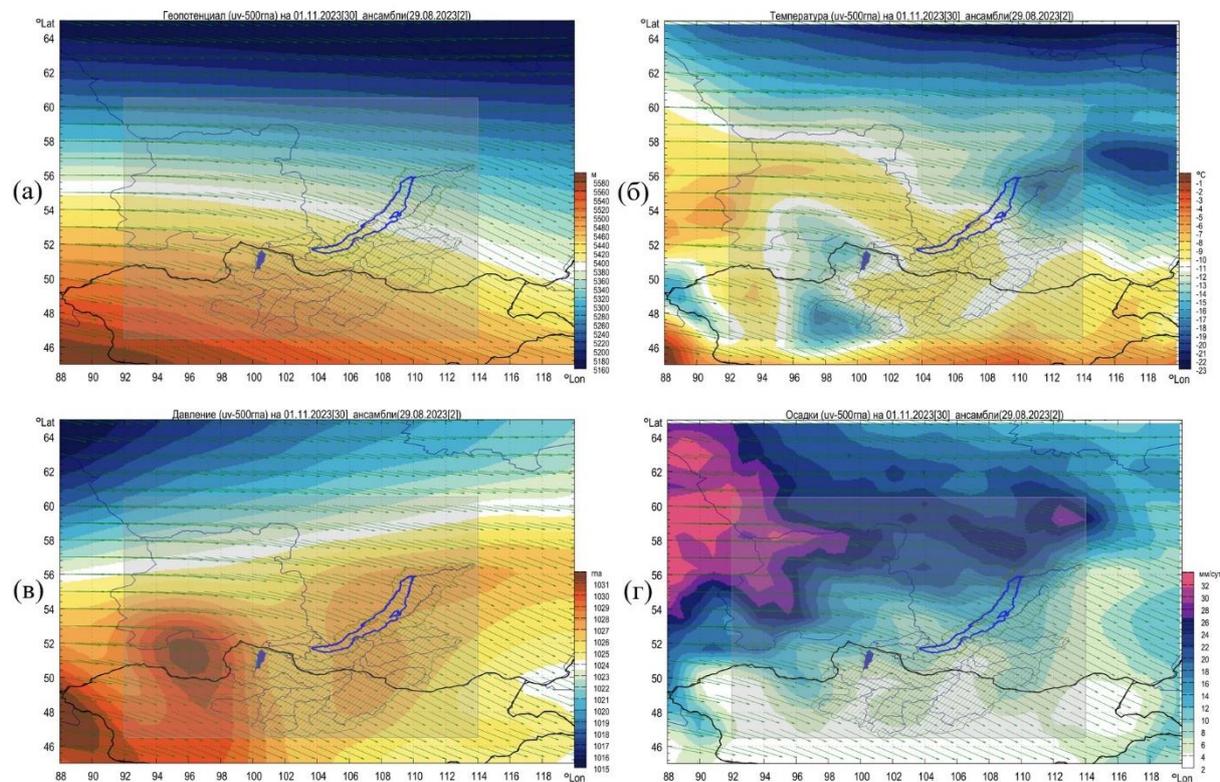


Рис. 7. Результаты визуализации прогностических данных в бассейне оз. Байкале за ноябрь 2023 г.

Прогностические показатели полезного притока в озеро позволяют проводить моделирование его уровня по различным диспетчерским графикам (правилам назначения расходов Иркутской ГЭС). Одним из важных по экологическим требованиям является диспетчерский график с максимальным приближением к естественным условиям регулирования (до строительства ГЭС).

Заключение. В ходе работы были рассмотрены форматы климатических данных; разработаны компоненты системы мониторинга и накопления фактических и прогностических глобальных данных по климату; разработаны компоненты визуализации этих данных.

Разработанная система мониторинга является портативной, позволяет оперативно обновлять данные с запаздыванием до нескольких суток и актуализировать их для отдельных локальных компьютеров и сетевых групп. Данная система позволяет заблаговременно получать оценки вероятных изменений климата, важных для эффективного управления как уровнем оз. Байкал, так и Ангарским каскадом ГЭС. В дальнейшем планируется расширение веб-интерфейса для анализа и формирования прогностических показателей.

Благодарности. Работа выполнена в рамках гранта № 075-15-2020-787 Министерства науки и высшего образования РФ на выполнение крупного научного проекта по приоритетным направлениям научно-технологического развития (проект «Фундаментальные

основы, методы и технологии цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории»).

Список источников

1. Абасов Н.В. Система долгосрочного прогнозирования и анализа природообусловленных факторов энергетики ГеоГИПСАР / Н.В. Абасов // Материалы международного совещания APN (MAIRS/NEESP/SIRS) Экстремальные проявления глобального изменения климата на территории Северной Азии (Enviromis-2012), 2012. – С. 63-66.
2. Abasov N.V., Nikitin V.M., Bereznykh T.V., Osipchuk E.N. Monitoring and predictive estimations of atmospheric parameters in the catchment area of lake Baikal. Atmosphere, 2022, vol. 13, iss. 1, article 49.
3. Abasov N.V., Nikitin V.M., Osipchuk E.N. A System of Models to study long-term operation of hydropower plants in the Angara cascade. Energy Systems Research, 2019, vol. 2, no. 2(6), pp. 5-18.
4. Berdnikov V.M., Abasov N.V. Prognostic maps of climatic indicators based on a multivariate neural network. Energy System Research, 2022, vol. 5, no. 4, pp. 19-23.
5. Petrukhina V.A., Abasov N.V. A technology for predictive estimation of meteorological parameters on the basis of the global climate model CFSv2. Energy System Research, 2022, vol. 5, no. 4, pp. 12-18.
6. Osipchuk E.N., Gasan V.S. Development of a Web Service for Analysis of the GeoGIPSAR Geoclimatic Data // Energy Systems Research, 2022, vol. 5, no. 3, pp. 5-12.
7. Набор данных ECMWF. – URL: <https://www.ecmwf.int/en/forecasts/datasets> (дата обращения: 10.03.2023).
8. Национальное управление океанических и атмосферных исследований. – URL: <https://www.noaa.gov/> (дата обращения: 10.03.2023).
9. Becker A., Finger P., Meyer-Christoffer A., et al. A description of the global land-surface precipitation data products of the global precipitation climatology centre with sample applications including centennial (trend) analysis from 1901–present. Earth Syst. Sci. Data, 2013, vol. 5, pp. 71-99.
10. Schamm K., Ziese M., Becker A., et al. Global gridded precipitation over land: A description of the new GPCP first guess daily product. Earth Syst. Sci. Data, 2014, vol. 6, pp. 49-60.
11. Schneider U., Finger P., Meyer-Christoffer A., et al. Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the global precipitation climatology centre (GPCP), Atmosphere, 2017, vol. 8, 52 p.
12. Оперативные данные сайта «Погода и климат». – URL: <http://www.pogodaiklimat.ru> (дата обращения: 10.03.2023).
13. Специальные данные для исследований климата. – URL: <http://aisori-m.meteo.ru/waisori/index0.xhtml> (дата обращения: March 10.03.2023).
14. Прогностические ансамбли. – URL: <https://nomads.ncep.noaa.gov/pub/data/nccf/com/cfs/> (дата обращения: 10.03.2023).
15. AIS GMVO. – URL: gmvo.skniivh.ru (дата обращения: 10.03.2023).
16. Данные о солнечной активности. – URL: <http://www.sidc.be/silso/DATA/> (дата обращения: 10.03.2023).
17. Руководство пользователя NetCDF – URL: <https://docs.unidata.ucar.edu/nug/current/index.html> (дата обращения: 10.03.2023).
18. Библиотека HDF5 и формат файлов – Группа HDF. – URL: <https://www.hdfgroup.org/solutions/hdf5> (дата обращения: 10.03.2023).
19. Центр прогнозирования климата. – URL: <https://www.cpc.ncep.noaa.gov/products/wesley/wgrib.html> (дата обращения: 10.03.2023).
20. Национальный центр атмосферных исследований. – URL: <https://ncar.ucar.edu/> (дата обращения: 10.03.2023).
21. Общественная зона Cosmo. – URL: <https://www.cosmo-model.org/> (дата обращения: 10.03.2023).
22. Saha S., Moorthi S., Wu X., et al. The NCEP climate forecast system version 2, J. Climate, 2014, no. 6(27), pp. 2185-2208.
23. Yuan X., Wood E., Luo L., et al. A first look at climate forecast system version 2 (CFSv2) for hydrological seasonal prediction, Geophys. Res. Lett., 2011, p. 38.
24. Rai A., Saha S.K. Evaluation of energy fluxes in the NCEP climate forecast system version 2.0 (CFSv2), Clim. Dyn., 2017, vol. 50, pp. 101-114.
25. Hourdin F., Mauritsen T., Gettelman A., et al. The Art and science of climate model tuning. Bulletin of the American Meteorological Society, 2017, vol. 98, pp. 589-602.

Гасан Виктор Святославович. Инженер-исследователь лаборатории гидроэнергетических и водохозяйственных систем Института систем энергетики им. Л.А. Мелентьева (ИСЭМ) СО РАН. Сфера научных интересов – веб-разработка и геоклиматическая аналитика, viktor-gasan@yandex.ru.

Абасов Николай Викторович. К.т.н., старший научный сотрудник ИСЭМ СО РАН. Сфера научных интересов – разработка новых методов и моделей для прогнозирования природообусловленных факторов энергетики и эффективного функционирования энергосистем, nva@iset.irk.ru.

Осипчук Евгений Николаевич. К.т.н., научный сотрудник ИСЭМ СО РАН. Сфера научных интересов – разработка моделей для оптимизации режимов работы ГЭС с учетом ограничений, критериев оптимизации и долгосрочных прогнозов притока, eugene.os@mail.ru.

UDC 004.94

DOI: 10.25729/ESI.2023.32.4.014

Monitoring and accumulation of data in the GeoGIPSAR information-forecasting system

Victor S. Gasan, Nikolai V. Abasov, Evgeny N. Osipchuk

Melentiev Energy Systems Institute (MESI) SB RAS, Irkutsk, Russia, *viktor-gasan@yandex.ru*

Abstract. The article discusses the source data, their structure and processing methods used in GeoGIPSAR information-forecasting system, which was developed at the Melentiev Energy Systems Institute (MESI) of SB RAS and continues to develop. It successfully user data from modern global climate model (such as NOAA, HPCC, etc.) to predict inflow at the large reservoirs and lakes with an assessment of their level modes. The article describes the structures and types of indicators of global climate models used in practice when forecasting runoff in river basins. The main steps for data collection, processing and analysis are considered. The results of the system operation using the developed map visualizer are demonstrated. It is noted that it is necessary to conduct research on forecasting a useful inflow for effective management of the level regime of Lake Baikal. On the results of the work carried out, it is planned to expand the system with the inclusion of a web interface for the analysis and formation of forecast indicators.

Keywords: global climate models, reanalysis, monitoring, information-forecasting system

Acknowledgements: The work was supported by the Ministry of Science and Higher Education of the Russian Federation, the grant No. 075-15-2020-787 for implementation of Major scientific projects on priority areas of scientific and technological development (the project “Fundamentals, methods and technologies for digital monitoring and forecasting of the environmental situation on the Baikal natural territory”).

References

1. Abasov N.V. Sistema dolgosrochnogo prognozirovaniya i analiza prirodo-obuslovlennykh faktorov jenergetiki GeoGIPSAR [The GeoGIPSAR system of long-term forecasting and analysis of environment-related factors of the energy industry], Materialy mezhdunarodnogo soveshhanija APN (MAIRS/NEESP/SIRS) “Jekstremal'nye proyavleniya global'nogo izmenenija klimata na territorii Severnoj Azii”: (Enviromis-2012) [Proc. of the International Meeting of APN: “Extreme manifestations of global climate change in North Asia”: Enviromis-2012], 2012, pp. 63-66.
2. Abasov N.V., Nikitin V.M., Berezhnykh T.V., Osipchuk E.N. Monitoring and predictive estimations of atmospheric parameters in the catchment area of lake Baikal. Atmosphere, 2022, vol. 13, iss. 1, article 49.
3. Abasov N.V., Nikitin V.M., Osipchuk E.N. A System of Models to study long-term operation of hydropower plants in the Angara cascade. Energy Systems Research, 2019, vol. 2, no. 2(6), pp. 5-18.
4. Berdnikov V.M., Abasov N.V. Prognostic maps of climatic indicators based on a multivariate neural network. Energy System Research, 2022, vol. 5, no. 4, pp. 19-23.
5. Petrukhina V.A., Abasov N.V. A technology for predictive estimation of meteorological parameters on the basis of the global climate model CFSv2. Energy System Research, 2022, vol. 5, no. 4, pp. 12-18.
6. Osipchuk E.N., Gasan V.S. Development of a Web Service for Analysis of the GeoGIPSAR Geoclimatic Data // Energy Systems Research, 2022, vol. 5, no. 3, pp. 5-12.
7. Dataset ECMWF, available at: <https://www.ecmwf.int/en/forecasts/datasets> (accessed: 04/10/2023).
8. National Oceanic and Atmospheric Administration available online <https://www.noaa.gov/> (accessed: 04/10/2023).

9. Becker A., Finger P., Meyer-Christoffer A., et al. A description of the global land-surface precipitation data products of the global precipitation climatology centre with sample applications including centennial (trend) analysis from 1901–present. *Earth Syst. Sci. Data*, 2013, vol. 5, pp. 71-99.
10. Schamm K., Ziese M., Becker A., et al. Global gridded precipitation over land: A description of the new GPCC first guess daily product. *Earth Syst. Sci. Data*, 2014, vol. 6, pp. 49-60.
11. Schneider U., Finger P., Meyer-Christoffer A., et al. Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the global precipitation climatology centre (GPCC), *Atmosphere*, 2017, vol. 8, 52 p.
12. Operational Data of the Site «Weather and Climate». Available online: <http://www.pogodaiklimat.ru> (accessed: 04/10/2023).
13. Dedicated Data for Climate Research. Available online: <http://aisori-m.meteo.ru/waisori/index0.xhtml> (accessed: 04/10/2023).
14. Forecasting ensembles, available at: <https://nomads.ncep.noaa.gov/pub/data/nccf/com/cfs/> (accessed: 04/10/2023).
15. AIS GMVO, available at: gmvo.skniivh.ru (accessed: 04/10/2023).
16. Solar activity data, available at: <http://www.sidc.be/silso/DATA/> (accessed: 04/10/2023).
17. NetCDF Users Guide, available at: <https://docs.unidata.ucar.edu/nug/current/index.html> (accessed: 04/10/2023).
18. The HDF5 Library & File Format – The HDF Group, available at: <https://www.hdfgroup.org/solutions/hdf5> (accessed: 04/10/2023).
19. Climate Prediction Center, available at: <https://www.cpc.ncep.noaa.gov/products/wesley/wgrib.html> (accessed: 04/10/2023).
20. National Center for Atmospheric Research, available at: <https://ncar.ucar.edu/> (accessed: 04/10/2023).
21. Cosmo Public area, available at: <https://www.cosmo-model.org/> (accessed: 04/10/2023).
22. Saha S., Moorthi S., Wu X., et al. The NCEP climate forecast system version 2, *J. Climate*, 2014, no. 6(27), pp. 2185-2208.
23. Yuan X., Wood E., Luo L., et al. first look at climate forecast system version 2 (CFSv2) for hydrological seasonal prediction, *Geophys. Res. Lett.*, 2011, p. 38.
24. Rai A., Saha S.K. Evaluation of energy fluxes in the NCEP climate forecast system version 2.0 (CFSv2). *Clim. Dyn.*, 2017, vol. 50, pp. 101-114.
25. Hourdin F., Mauritsen T., Gettelman A., et al. The Art and Science of Climate Model Tuning, *Bull. Am. Meteorol. Soc.*, 2017, vol. 98, pp. 589-602.

Viktor Gasan Svyatoslavovich. *Research engineer of the Laboratory of Hydropower and Water Management Systems of Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences (MESI SB RAS). His research interests include web-development and the geoclimatic analytics, viktor-gasan@yandex.ru.*

Abasov Nikolai Viktorovich. *Ph.D. in technology, senior researcher at MESI SB RAS. His research interests include development of new methods and models for predicting environmental factors in the energy sector and the efficient functioning of energy systems, nva@isem.irk.ru.*

Osipchuk Evgeny Nikolaevich. *Ph.D. in technology, researcher at MESI SB RAS. His research interests include development of new models for optimizing HPP regimes, taking into account constraints, optimization criteria and long-term inflow forecasts, eugene.os@mail.ru.*

Статья поступила в редакцию 01.09.2023; одобрена после рецензирования 01.12.2023; принята к публикации 16.12.23.

The article was submitted 09/01/2023; approved after reviewing 12/01/2023; accepted for publication 12/01/2023.

УДК 004.738

DOI: 10.25729/ESI.2023.32.4.015

Выделение групп используемых Интернет-ресурсов для обнаружения внутренних источников киберугроз

Исаев Сергей Владиславович, Донцов Денис Юрьевич

Институт вычислительного моделирования СО РАН,

Россия, Красноярск, *si@icm.krasn.ru*

Аннотация. Защищенность корпоративной сети является важным аспектом успешного функционирования организации. В данной работе исследуется безопасность внутреннего периметра сети на примере Красноярского научного центра СО РАН. Существуют различные средства для предотвращения киберугроз и анализа посещаемых Интернет-ресурсов, но их быстродействие и возможность применения сильно зависят от объема входных данных. В статье рассматриваются существующие методы определения сетевых угроз с помощью анализа журналов прокси-сервера. Исследовано разделение интернет-пользователей на тематические группы для выявления аномалий. Предложен метод кластеризации Интернет-ресурсов, направленный на снижение объема входных данных путем исключения групп безопасных Интернет-ресурсов или выбора только подозрительных Интернет-ресурсов. Предложенный метод состоит из этапов: предобработка данных, выделение сессий пользователей, анализ данных и интерпретация полученных результатов. Исходными данными являются записи журнала прокси-сервера. На первом этапе из исходных данных выбираются полезные для анализа данные, после чего непрерывный поток данных делится на небольшие порции (сессии) при помощи метода ядерной оценки плотности. На втором этапе выполняется мягкая кластеризация используемых Интернет-ресурсов путем применения метода тематического моделирования. Результатом второго этапа являются неразмеченные группы Интернет-ресурсов. На третьем этапе, с помощью эксперта, происходит интерпретация полученных результатов путем анализа наиболее популярных Интернет-ресурсов в каждой группе. Метод имеет множество настроек на каждом этапе, что позволяет сконфигурировать его под любой формат и специфику входных данных. Область применения метода не ограничена. Он может быть использован как в качестве дополнительного шага предобработки с целью снижения количества входных данных, так и при выявлении аномальных данных.

Ключевые слова: кластерный анализ, тематическое-моделирование, кибербезопасность

Цитирование: Исаев С.В. Выделение групп используемых Интернет-ресурсов для обнаружения внутренних источников киберугроз / С.В. Исаев, Д.Ю. Донцов // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 170-181. – DOI:10.25729/ESI.2023.32.4.015.

Введение. Повсеместное внедрение информационно-телекоммуникационных технологий в жизнь людей и переход на цифровую экономику требуют повышенного внимания к проблемам кибербезопасности. Выполнение профессиональных обязанностей в настоящее время в большой мере связано с использованием информационно-телекоммуникационных ресурсов. Современные текстовые редакторы, установленные непосредственно на компьютере, требуют подключения к сети Интернет для загрузки справочных материалов, шаблонов оформления или обновления программных модулей. Даже в быту уже сложно себе представить жизнь без различных электронных устройств, облегчающих процесс получения информации и обеспечивающих коммуникации между людьми. Во многих производственных процессах используется сеть Интернет для обеспечения сервисного обслуживания и удаленного контроля состояния оборудования. Повсеместное внедрение технологии «интернета вещей» повышает риск несанкционированного доступа к цифровым активам за счет недостаточной защиты или некорректных настроек отдельных устройств. В связи с массовостью явления невозможно в каждом случае привлечь квалифицированного эксперта в области кибербезопасности для выявления и предотвращения возможных рисков. Необходимостью становятся средства защиты, ко-

которые не требуют специальных знаний по настройке и способны противодействовать изменяющимся угрозам. Этим критериям соответствуют антивирусы, встроенные средства защиты, сканеры безопасности, системы предотвращения вторжений и прочие средства защиты, которые получают периодические обновления от разработчиков. Они решают множество задач, но, к сожалению, не охватывают весь спектр проблем кибербезопасности.

Можно выделить три основных класса источников киберугроз: связанные с действиями человека, проблемы функционирования техники и связанные с форс-мажорными ситуациями [1]. В большинстве случаев именно человек является причиной возникновения киберугроз [2], поэтому перспективным направлением является создание методов и алгоритмов, помогающих снизить риск вторжений по вине человека. Чтобы блокировать доступ к потенциально опасным Интернет-ресурсам, многие крупные организации используют их фильтрацию [3]. Это позволяет существенно снизить риск киберугроз, но может создать проблемы доступа к нужным ресурсам и не дает 100% защиту. В связи с этим, создание дополнительных средств защиты, учитывающих особенности и сценарии использования Интернет в конкретной организации, является актуальным.

Анализ кибербезопасности с точки зрения внутреннего периметра сети требует получения и сохранения в пригодном для обработки виде запросов пользователей и устройств сети, метаданных о структуре сети и т.д. В результате анализа можно определить аномалии сценариев взаимодействия в сети, идентифицировать и оперативно блокировать угрозы [4-6]. В процессе работы телекоммуникационных программ и протоколов образуется большое число всевозможных журналов функционирования сервисов и использования Интернет-ресурсов. Они располагаются как на устройствах пользователей, так и на телекоммуникационном оборудовании. Суммарный объем этих данных для крупной организации за один день составляет десятки Гигабайт. В большинстве работ используются методы анализа, основанные на кластеризации [7, 8] или машинном обучении [9, 10]. Выделение части данных, полезных для анализа безопасности, позволяет снизить время их обработки, увеличить временной диапазон для анализа, что позволяет повысить оперативность и качество решений по защите корпоративной сети. Уменьшить объем сохраняемых для последующего анализа данных можно за счет группировки посещаемых пользователями ресурсов и определения среди них условно безопасных и потенциально опасных. Кроме того, подобная группировка полезна с точки зрения обнаружения аномальных групп ресурсов, сигнализирующих о потенциальных киберугрозах. Множество имеющихся коммерческих сетевых анализаторов (Kaspersky Anti Targeted Attack, PT Network Attack Discovery, Threat Detection System и др.) нацелены в первую очередь на обнаружение и предотвращение атак из внешнего периметра.

В данной работе исследуется возможность анализа данных запросов пользователей к ресурсам сети Интернет с целью идентификации рисков кибербезопасности и сокращения исходного набора данных. Предложена методика мягкой кластеризации посещаемых пользователями Интернет-ресурсов на основе применения тематического моделирования данных журналов доступа прокси-сервера. Основная цель состоит в исключении из дальнейшего анализа найденных групп с безопасным контентом и выделении аномальных групп, несущих повышенные риски кибербезопасности.

Источники данных. В данном исследовании за основу были взяты данные ежедневных журналов прокси-сервера корпоративной сети Красноярского научного центра СО РАН. Сбором и анализом данных занимается отдел информационно-телекоммуникационных технологий Института вычислительного моделирования СО РАН, который более двух десятков лет обеспечивает доступ пользователей к различным Интернет-сервисам. За это время были

накоплены значительные объемы данных, проводились исследования различных аспектов информационной безопасности корпоративной сети научного центра [11]. На основе анализа данных были выявлены источники киберугроз и предложены эффективные способы их блокирования [12]. В большинстве исследований анализируются лишь внешние угрозы кибербезопасности, источники которых расположены в неконтролируемой зоне сети Интернет, но угрозы безопасности внутри корпоративной сети тоже несут значительные риски и не имеют единой точки защиты. К таким угрозам можно отнести: вирусы, попадающие во внутреннюю сеть на флеш-накопителях и дисках, мобильные компьютеры и электронные устройства, бесконтрольно подключаемые ко внутренней сети. Если на таком устройстве имеется вирус или некорректно настроенная программа, то обнаружить их можно по нетипичным действиям в сети, в том числе попыткам получить доступ в Интернет. Так как рассматриваемые устройства и находящиеся на них программы не используют штатные для сети настройки, то их возможно идентифицировать и заблокировать путем анализа сетевых запросов и выявления аномалий.

Одним из множества источников данных об использовании ресурсов сети Интернет работниками организации может служить журнал прокси-сервера, который выступает посредником между веб-браузером и веб-сервером или другим Интернет-сервисом. Журналы прокси-сервера содержат данные о доступе пользователей к сети Интернет. На основе их анализа можно решать задачи оптимизации телекоммуникационной подсистемы, улучшения защиты, выявления нетипичных запросов и аномальной активности устройств.

Рассматриваются данные ежедневных журналов прокси-сервера за месяц, общий объем которых составил более 2 Гигабайт. Суммарное количество сохраненных запросов к Интернет-ресурсам превышает десять миллионов, что выводит используемые для анализа алгоритмы в область Big-data. Журнал представляет собой текстовый файл, в каждой строке которого содержатся следующие атрибуты:

- время запроса по часам прокси-сервера;
- длительность отклика;
- IP-адрес клиента источника запроса;
- статус запроса прокси сервера;
- HTTP-код состояния, отправленный клиенту;
- размер переданных клиенту данных;
- метод совершения запроса;
- унифицированный адрес ресурса;
- идентификатор пользователя;
- тип содержимого запроса по MIME.

На основе группировки строк по идентификатору пользователя и IP-адресу клиента можно для каждого источника запросов вычислить дополнительные метаданные, такие, как: время начала и конца активности, продолжительность активности, средняя частота и размер полученных данных.

Выделение групп пользователей. Целью анализа являлось выделение тематических групп пользователей Интернет-ресурсов. Предпосылкой появления угроз кибербезопасности в данной постановке является изменение тематической группы пользователем или наличие аномально малых групп. Последовательность действий для анализа представлена на схеме (рисунок 1).

Для дальнейшей обработки использовались записи об успешно полученных ресурсах, имеющие код 200, так как в случае недоступного ресурса невозможно судить о его содержании и, в дальнейшем, проверить правильность классификации. Из них были выбраны параметры с

унифицированным адресом ресурса (URL) и идентификатором пользователя (UserName). Были созданы: список пользователей, список ресурсов и таблица, связывающая пользователей, ресурсы и количество посещений. После обработки журналов за семь дней объем данных уменьшился примерно на три порядка с 2 ГБ до 3 МБ. Полученный таким образом набор данных для одного пользователя может быть интерпретирован как точка в многомерном пространстве, где оси являются ресурсами, а количество посещений ресурса – координатой точки на оси. Для не посещенных ресурсов некоторым пользователем (отсутствующие данные) принимается значение координаты равным 0.

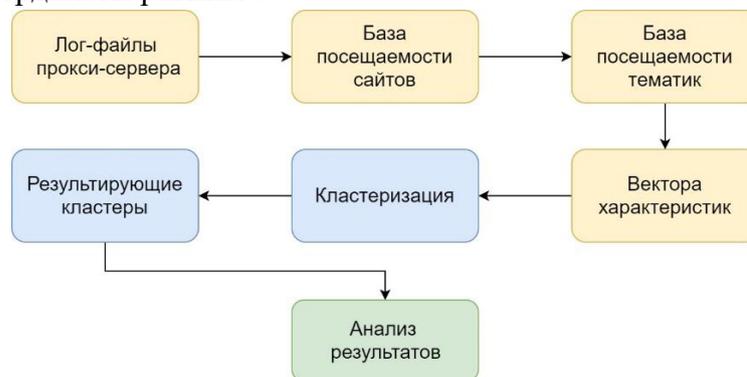


Рис. 1. Последовательность выделения групп пользователей

Ввиду того, что ресурсов, имеющих различный URL, очень много, были просуммированы ресурсы с одинаковым доменным именем, что позволило снизить размерность пространства до 4000. Для решения задачи дальнейшего понижения количества показателей было решено произвести свертку и перейти от доменных имен Интернет-ресурсов к тематикам сайтов. Помимо снижения размерности, это позволяет получить для каждого пользователя его предпочтительные тематики Интернет-ресурсов.

Чтобы решить задачу определения тематики сайтов, был использован сервис Яндекс-Каталог (<https://yandex.ru/yaca>), позволяющий определить тематическую рубрику для большинства сайтов. К моменту начала исследований это был наиболее полный и удобный для автоматизированного использования сервис. К сожалению, тематики некоторых не популярных или малопосещаемых сайтов не определяются данным сервисом, поэтому их пришлось исключить из дальнейшего анализа, при этом какие-то полезные для решения задачи данные были потеряны, но была достигнута полная автоматизация процесса. В результате, получившийся набор данных имеет порядка 160 атрибутов, что на три десятичных порядка меньше первичного количества различных Интернет-ресурсов. Для такой размерности могут эффективно применяться различные методы кластеризации. Таблица 1 содержит пример фрагмента данных с количеством посещений по тематикам.

Таблица 1. Число посещений по тематикам

Идентификатор пользователя	Тематики из «Яндекс.Каталог»			
	Интернет	Социальные сети	Баннерные сети	Банки
4	3102	58	0	0
7	227	8	73	8
5	142	2	0	0
6	113	14	226	56
2	80	1	125	0
3	78	21	21	15
1	24	7	0	0

Для применения методов кластерного анализа потребовалась нормализация всех характеристик к единому диапазону. После нормализации всех характеристик эти данные можно представить в виде многомерного пространства, в котором каждая координата – это значение показателя. При кластеризации существенным является способ задания расстояния между отдельными кластерами. Были рассмотрены несколько алгоритмов вычисления расстояния, в результате наиболее подходящим был признан метод минимальной дисперсии Уорда. Расстояние вычисляется по формуле:

$$d(U, V) = \sqrt{\frac{|V|+|S|}{Q} * d(V, S) + \frac{|V|+|T|}{Q} * d(V, T) - \frac{|V|}{T} * d(S, T)^2},$$

где кластер U был получен путем объединения кластеров T и S , а $Q = |V| + |T| + |S|$, под операцией $|*|$ понимается нахождение количества элементов в соответствующем кластере.

Для получения групп применялся метод иерархической кластеризации (рисунок 2).

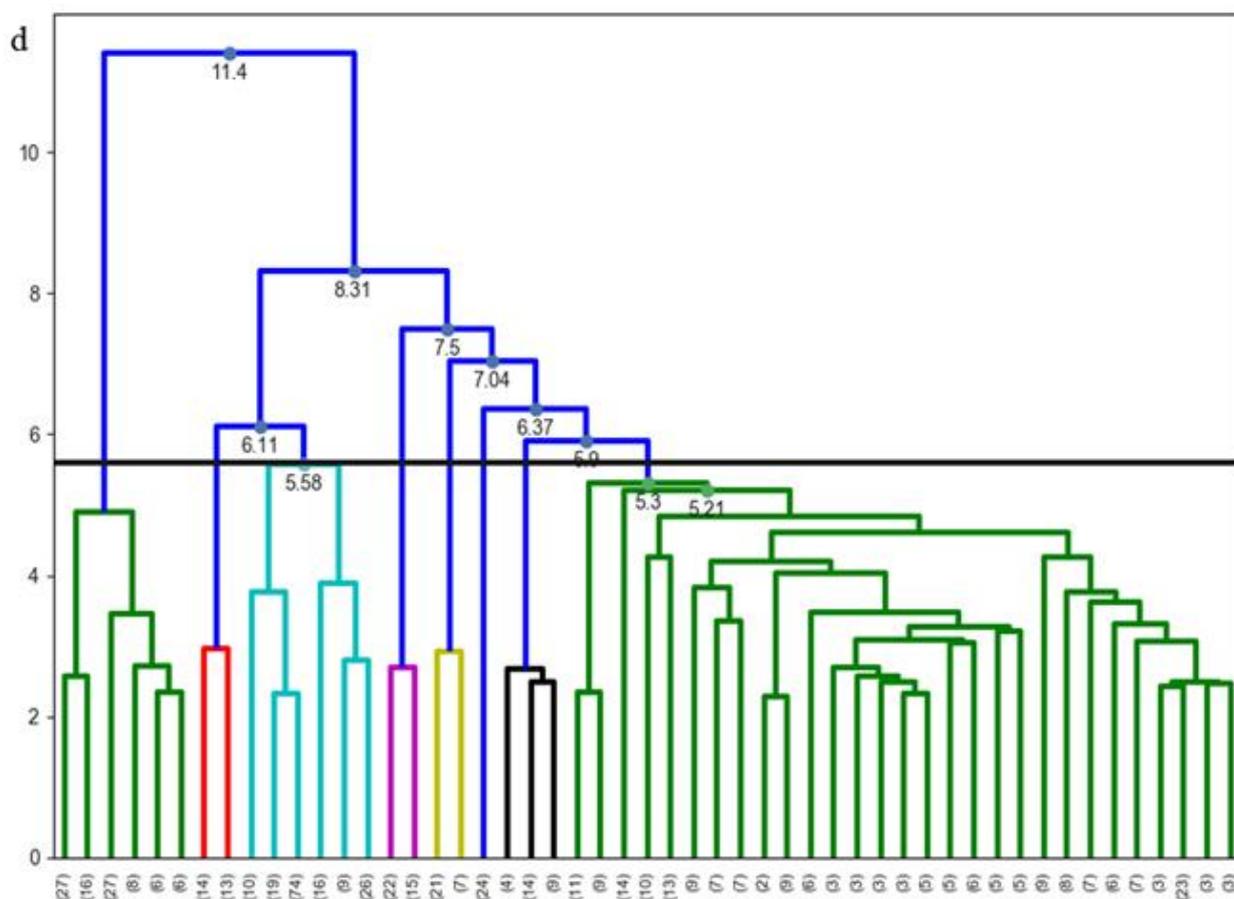


Рис. 2. Дендрограмма иерархической кластеризации источников

Для визуализации полученных групп были использованы методы главных компонент (PCA) и линейного дискриминантного анализа (LDA). Для количества кластеров 3 метод LDA показал более наглядные результаты (рисунок 3), при большем количестве кластеров оба метода не позволяют разграничить кластеры на плоскости.

Для проверки гипотезы о постоянстве предпочтений пользователей произведено сравнение полученных групп за разные промежутки времени. Это необходимо для подтверждения того, что пользователи из одного кластера сохраняют схожие предпочтения за определенный период времени. Смена предпочтений при этом сигнализировала бы о аномалии. В результате анализа групп за несколько интервалов времени был сделан вывод о наличии постоянных

групп, в которые входят более половины пользователей. Таким образом, признаком возможной угрозы при таком способе группировки остается только появление групп из единичных пользователей при небольшой глубине кластеризации. Кроме того, охват посещаемых сайтов сервиса Яндекс-Каталог недостаточен, в связи с чем существенная часть ресурсов выпала из рассмотрения.



Рис. 3. Результат визуализации 3-х кластеров методами PCA и LDA

Выделение групп Интернет-ресурсов. Для устранения зависимости от внешних классификаторов и включения в рассмотрение всех Интернет-ресурсов был исследован подход по установлению зависимостей между посещаемыми Интернет-ресурсами в пределах некоторых сессий и получению групп Интернет-ресурсов, имеющих зависимости в различных сессиях. Сессией будем называть совокупность последовательно посещенных пользователем Интернет-ресурсов за некоторый период активности. В простейшем случае в качестве сессии можно выбирать одни сутки, но очевидно, что за сутки у пользователя как правило есть несколько сеансов работы с Интернет-ресурсами, и в разные сеансы могут решаться различные задачи. Для улучшения качества кластеризации нужно использовать и прочие способы определения сессий.

Для решения задачи разделения Интернет-ресурсов на группы путем анализа их принадлежности одной сессии будем использовать вероятностное тематическое моделирование [13]. Вероятностные тематические модели осуществляют «мягкую» кластеризацию, позволяя документу или термину относиться сразу к нескольким темам с различными вероятностями. Документом в нашем случае является множество Интернет-ресурсов, посещенных пользователем в течение сессии. Термином является сам Интернет-ресурс, независимо от времени и пользователя, его посетившего. В результате моделирования мы получаем наборы Интернет-ресурсов, которые совместно встречались в разных сессиях. Каждому такому набору можно подобрать условную тематику. Например, если в исходном журнале присутствуют записи:

$$R_{1,1}, R_{2,1}, R_{3,1}, R_{4,2}, R_{5,2}, R_{6,1}, R_{7,1}, R_{8,2}, R_{9,2}, R_{10,1},$$

где первый индекс – номер записи, а второй – номер пользователя, то после разбиения на сессии мы можем получить:

- $S_{1,1} = (R_{1,1}, R_{2,1}, R_{3,1})$ – первая сессия первого пользователя;
- $S_{1,2} = (R_{4,2}, R_{5,2}, R_{8,2}, R_{9,2})$ – первая сессия второго пользователя;
- $S_{2,2} = (R_{6,1}, R_{7,1}, R_{10,1})$ – вторая сессия первого пользователя.

Результатом тематического моделирования является набор неименованных групп (тематик), включающих расположенные в порядке количества встречаемости Интернет-ресурсы: $T_1 = (R_{1,1}, R_{1,2}, R_{1,3}, \dots)$, $T_2 = (R_{2,1}, R_{2,2}, R_{2,3}, \dots)$, $T_3 = (R_{3,1}, R_{3,2}, R_{3,3}, \dots)$, ...

Эксперт может на основании содержания наиболее популярных Интернет-ресурсов группы присвоить ей тематику и определить степень опасности. Безопасные группы мы можем удалить из дальнейшего анализа, а самые опасные исследовать более детально. Задача решалась в несколько этапов.

Этап 1 – Предобработка данных. Современные веб-сайты устроены таким образом, что, открывая один адрес, браузер запрашивает, помимо основной ссылки, еще 10-20 других, в которых могут запрашиваться как необходимые для отображения страницы элементы, так и не имеющие визуального представления счетчики и элементы программного кода. В среднем за рабочий день в журнале прокси-сервера регистрируется около четырех тысяч запросов от каждого пользователя. Целью предобработки является удаление из набора данных, не несущих полезной для дальнейшего анализа информации, за счет чего удается повысить качество результатов и скорость основной фазы анализа [14].

Из имеющихся данных журналов были исключены записи, касающиеся анонимного доступа для разрешенного технического трафика (обновления известных приложений и сервисов) и ресурсы типа `css/js/image`, которые, как правило, не несут смысловой нагрузки или ее сложно определить автоматически без привлечения эксперта. За счет этого объем данных был снижен примерно в 5 раз. Включение в анализ ресурсов JavaScript требует отдельного рассмотрения и лежит вне пределов данной работы.

Следующим шагом предобработки было выделение из URL частей, содержащих полный домен или IP-адрес использованного ресурса, для суммирования числа посещений ресурсов с одинаковой доменной частью.

Этап 2 – Выделение сессий пользователей. Требуется сгруппировать данные каждого пользователя в сессии с непересекающимися временными интервалами. Имеется несколько возможных алгоритмов получения таких наборов, отличающихся сложностью и достоверностью полученных данных.

1. Наборы с постоянным периодом. Для выделения таких сессий нужно задать постоянный временной интервал и выделить данные пользователя внутри него. При использовании такого подхода в одну сессию могут попадать периоды активности пользователя небольшой длительности. Например, утром пользователь работал с почтовой программой, а вечером того же дня читал новости. При длине периода 24 часа это будет считаться одной сессией, что не согласуется с реальностью.
2. Определение периодов отсутствия активности пользователя и выделений сессий между этими периодами. Такой подход генерирует сессии произвольной длины, в зависимости от определения минимального периода неактивности. Существенным недостатком метода является его неработоспособность при наличии постоянных фоновых процессов, таких, как работа мессенджера, почтового клиента и прочих программ, периодически опрашивающих интернет-ресурс, независимо от активности пользователя.

Для устранения проблем перечисленных подходов предложено использовать метод ядерной оценки плотности (Kernel Density Estimation – KDE), основанный на непараметрической оценке плотности случайной величины [15-17]. Этот метод оценивает плотность распределения одномерного набора данных и обнаруживает точки локального экстремума. Используя, например, точки минимума в качестве границ интервалов активности, мы получаем сессии различной длины, соответствующие реальной активности пользователя. На исследуемых данных средняя длина выделенной сессии составила 4.5 минуты.

Для применения данного метода необходимо задать два параметра – размер ядра (в нашем случае это количество ресурсов, посещенных пользователем за сессию), и ширина полосы (диапазон времени для агрегирования). Изменение этих параметров существенно влияет

на количество выделенных сессий. Их значения подбирались опытным путем, на основе экспертной оценки достоверности размеров получаемых сессий. На рисунке 4 показана гистограмма распределения среднего размера ядра среди всех пользователей. По горизонтальной оси размер ядра, по вертикальной – количество пользователей.

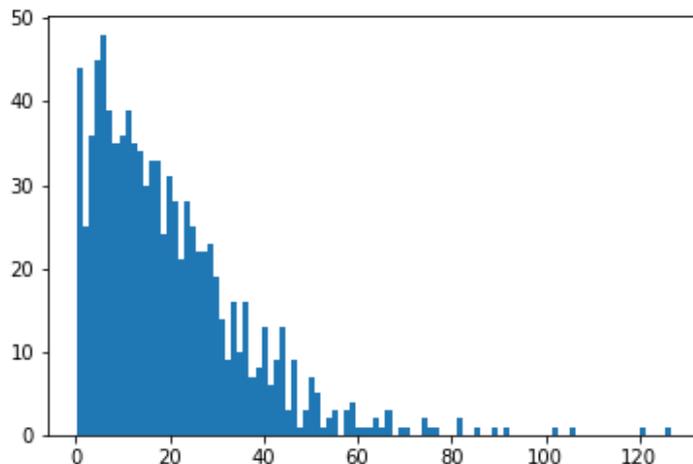


Рис. 4. Гистограмма распределения среднего размера ядра

Этап 3 – Тематическое моделирование. Решив задачу разделения данных журналов на сессии, мы получили все начальные данные для применения тематического моделирования для мягкой кластеризации наших сессий, состоящих из Интернет-ресурсов. В настоящее время известны и используются различные методы тематического моделирования, такие, как: методы сингулярного разложения (SVD), метод моментов, вероятностный латентно-семантический анализ (PLSA) и др. [18-19]. В данной работе использовался наиболее распространённый и хорошо описанный метод латентного размещения Дирихле (LDA) [20-21]. Применение метода LDA для полученных сессий позволяет мягко сгруппировать все Интернет-ресурсы на заданное количество групп, которое должно определяться опытным путем, в зависимости от ожидаемого результата. В таблице 2 приведен результат тематического моделирования ресурсов, посещенных пользователями за один месяц. Ресурсы разделены на 5 групп и расположены в порядке убывания силы принадлежности к группе.

Таблица 2. Результат тематического моделирования для 5 групп

№ группы	Ресурсы в порядке убывания принадлежности группе
1	newslab.ru, 4pda.ru, sfkras.ru, edu.sfu-kras.ru, worldcrisis.ru, libgen.is
2	nowa.cc, ugadalki.ru, scask.ru, forum.rcmir.com, 2baksa.net, autoopt.ru
3	update.eset.com, law-college-sfu.ru, kinoaction.ru, kiwt.ru, dostavka-krasnoyarsk.ru
4	apps.webofknowledge.com, packages.linuxmint.com, http.debian.net, urod.ru, fips.ru, mc.corel.com
5	fitohobby.ru, ib.adnxs.com, allrefs.net, ckp-rf.ru, teammodels.no, profinance.ru

На рисунке 5 представлен результат тематического моделирования с числом групп 30, с проекцией на две главные компоненты. Размер окружности зависит от количества посещений элементов группы.

В результате проведенного тематического моделирования мы выделили потенциально опасные группы с малым количеством посещений и расположенные обособленно от прочих. В большинстве случаев это оказались ресурсы, связанные со специализированными задачами

по загрузке или синхронизации информации через Интернет. Получение информации о таких задачах позволило их упорядочить и регламентировать.

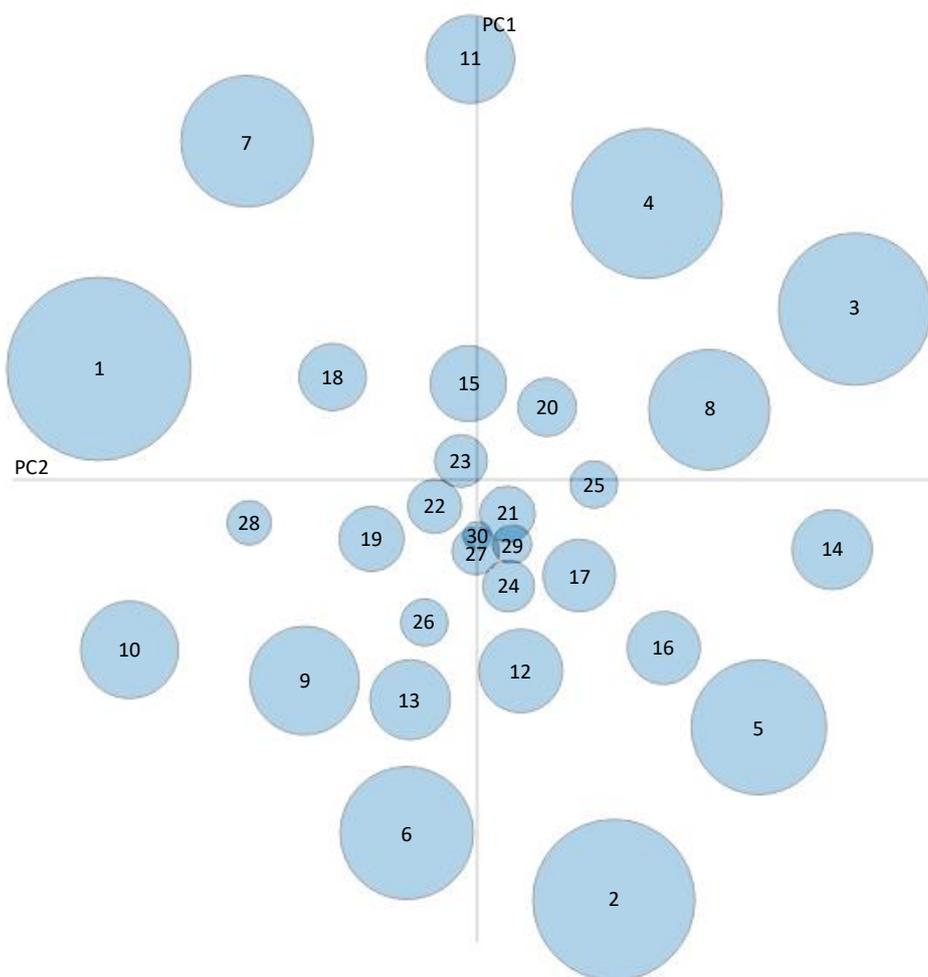


Рис. 5. Визуализация результатов тематического моделирования

Заключение. В работе рассмотрены два способа разбиения Интернет-ресурсов: с помощью внешнего классификатора и методом тематического моделирования. В первом случае в результате кластеризации мы получаем тематические группы пользователей с явно прописанными тематиками. На основании этого выявляются пользователи с аномальным поведением или посещающие ресурсы с потенциально вредоносной тематикой. К сожалению, использование внешнего классификатора накладывает существенные ограничения на возможности автоматического анализа неизвестных классификатору ресурсов, за счет чего возможно пропустить источники рисков. Перспективным представляется использование нескольких сервисов-классификаторов, которые бы покрывали большее множество Интернет-ресурсов, но даже в таком случае новые ресурсы будут в них отсутствовать.

Второй подход, выполняющий разбиение данных на сессии и тематическое моделирование, использует для анализа весь объем данных. Наличие настраиваемых параметров позволяет получать результаты как для организации с малым количеством пользователей и слабым каналом Интернет, так и для высоконагруженной корпоративной сети. Использование подхода возможно как на сетях без существующих аномалий, так и с функционирующими аномальными источниками, так как мы в любом случае получим группы ресурсов, которые должны будем классифицировать автоматически или с привлечением экспертов. Благодаря

мягкому разбиению Интернет ресурсов на тематики можно решать задачи: определение тематик Интернет-ресурсов, выявление сферы интересов пользователей, выделение сайтов с безопасной и несущей риски тематикой, а также определение и блокирование сервисов Интернет-рекламы.

Для улучшения результатов моделирования в дальнейшем планируется использование различного рода метаданных, таких, как длительность сессии, преимущественные виды контента, частота запросов. Описанный подход может быть использован для выявления внешних источников угроз за счет анализа журналов доступа к веб-ресурсам организации. Кластеризация источников с помощью тематического моделирования может помочь в выявлении источников угроз, которые не детектируются обычными системами обнаружения вторжений в связи с малыми временными рамками анализа.

Визуализация полученных групп с проекцией на главные компоненты позволяет получить общее представление о тематиках Интернет-запросов пользователей и, в отдельных случаях, выявить аномалии. Разработанные программные компоненты могут быть использованы при создании инструментальной системы обработки и анализа данных журналов Интернет-сервисов, а предложенные решения по противодействию обнаруженным угрозам встраиваться в функционирующие системы информационной безопасности.

Список источников

1. Mouna J., Latifa B., Latifa B.R., Anis A. Classification of security threats in information systems. *Procedia Computer science*, 2014, vol. 32, p. 489-496.
2. Дерендяев Д.А. Определение влияния человеческого фактора на основные характеристики угроз безопасности / Д.А. Дерендяев, Ю.А. Гатчин, В.А. Безруков // *Кибернетика и программирование*, 2019. – № 3. – С. 38-42.
3. Gyrodi R., Cornelia G., Pecherle G., Radu L. Network security using firewalls. *Journal of Computer science and control systems*, 2008, vol. 1.
4. Kao D.Y., Wang S.J., Huang F. Dataset Analysis of proxy logs detecting to curb propagations in network attacks. *Intelligence and security informatics*, 2008, pp. 245-250.
5. Marshall B., Chen H. Using importance flooding to identify interesting networks of criminal activity. *Lecture notes in computer science*, 2006, vol. 3975, pp. 14-25.
6. Mulkamala S., Sung A. Identifying significant features for network forensic analysis using artificial techniques. *International journal of Digital evidence*, 2003, vol. 1, no 4.
7. Bayraktar C., Karakaya Z., Gökçen H. Real-time anomaly detection system within the scope of smart factories. *The journal of Supercomputing*, 2023, vol. 79.
8. Wang C., Zhou H., Hao Z., Hu S., Li J., Zhang X., Jiang B., Chen X., Network traffic analysis over clustering-based collective anomaly detection. *Computer networks*, 2022, vol. 205, pp. 108760.
9. Yang T., Jiang Z., Liu P., Yang Q., Wang W. A traffic anomaly detection approach based on unsupervised learning for industrial cyber-physical system. *Knowledge-Based Systems*, 2023, vol. 279, pp. 110949.
10. Demertzis K., Tsiknas K., Taketzis D., Skianis C., Iliadis L. Darknet Traffic big-data analysis and network management for real-time automating of the malicious intent detection process by a weight agnostic neural networks framework. *Electronics*, 2021, vol. 10, pp. 781.
11. Исаев С.В. Кибербезопасность научного учреждения - активы и угрозы / С.В. Исаев // *Информатизация и связь*, 2015. – №1. – С. 53-57.
12. Исаев С.В. Анализ киберугроз и их источников для корпоративной сети Красноярского научного центра СО РАН / С.В. Исаев / *Информационные и математические технологии в науке и управлении*, 2016. – № 4-1. – С. 76-85.
13. Blei D.M. Probabilistic topic models. *Communications of the ACM*, 2012, vol. 55, no. 4, pp. 77-84.
14. Fei, B., Eloff, J., Oliver, M., Venter, H. Analysis of web proxy logs. *IFIP international conference on digital forensics*, Orlando, 2006, vol. 222, pp. 247-258.
15. Scott D.W. *Multivariate density estimation. Theory. Practice and visualization: second edition*. New York, 2015.
16. King T.L., Bentley R.J., Thornton L.E. [et al.] Using kernel density estimation to understand the influence of neighbourhood destinations on BMI. *BMJ Open*, 2016, vol. 6.

17. Kalinic M., Krisp J. Kernel density estimation (KDE) vs. hot-spot analysis - Detecting criminal hot spots in the city of San Francisco. Lund, Sweden, 2018.
18. Воронцов К.В. Обзор вероятностных тематических моделей, 2021.
19. Albalawi R., Yeap T., Benyoucef M. Using topic modeling methods for short-text data: A comparative analysis. Frontiers in artificial intelligence, 2020, vol. 3.
20. Jelodar H., Wang Y., Yuan, Ch., Xia, F. Latent dirichlet allocation (LDA) and topic modeling: models, applications, a survey, 2017.
21. Tharwat A., Gaber T., Ibrahim A., Hassanien A.E. Linear discriminant analysis: A detailed tutorial. Ai communications, 2017, vol. 30, pp. 169-190.

Исаев Сергей Владиславович. Кандидат технических наук, доцент, Институт вычислительного моделирования СО РАН, заместитель директора по научной работе, зав. отделом информационно-телекоммуникационных технологий. Область научных интересов: кибербезопасность и защита информации, интернет-технологии, распределённые информационные системы, интеллектуальные системы. AuthorID: 1623, SPIN: 1552-8542, si@icm.krasn.ru, Россия, 660036, Красноярск, Академгородок, дом. 50, стр. 44, ИВМ СО РАН.

Донцов Денис Юрьевич. Аспирант, Институт вычислительного моделирования СО РАН. Область научных интересов: машинное обучение, интернет-технологии. AuthorID: 1623, doncov.dy@icm.krasn.ru, Россия, 660036, Красноярск, Академгородок, дом. 50, стр. 44, ИВМ СО РАН.

UDC 004.738

DOI: 10.25729/ESI.2023.32.4.015

Identification of groups of visited Internet resources for detection of internal cyberthreats source

Sergey V. Isaev, Denis Y. Doncov

Institute of Computational Modeling SB RAS, Russia, Krasnoyarsk, si@icm.krasn.ru

Abstract. The protection of the corporate network is an important aspect of the successful functioning of the organization. In this paper, the cybersecurity of the internal network perimeter is studied using the example of the Krasnoyarsk Scientific Center of the Siberian Branch of the Russian Academy of Sciences. There are various tools for preventing cyber threats and analyzing visited Internet resources, but their performance and applicability strongly depend on the amount of input data. The article discusses existing methods for identifying network threats by analyzing proxy server logs. The division of Internet users into thematic groups to detect anomalies is investigated. A method for clustering Internet resources is proposed, aimed at reducing the volume of input data by excluding groups of safe Internet resources or selecting only suspicious Internet resources. The proposed method consists of the following steps: data preprocessing, user session selection, data analysis, and interpretation of the results. The source data is the log entries of the proxy server. At the first stage, useful data for analysis are selected from the initial data, after which the continuous data stream is divided into small portions (sessions) using the kernel density estimation method. At the second stage, soft clustering of the used Internet resources is performed by applying the topic modeling method. The result of the second stage are unallocated groups of Internet resources. At the third stage, with the help of an expert, the results obtained are interpreted by analyzing the most popular Internet resources in each group. The method has many settings at each stage, which allows you to configure it for any format and specifics of the input data. The scope of the method is not limited. It can be used both as an additional preprocessing step to reduce the amount of input data and to detect anomalous data.

Keywords: cluster analysis, topic-modeling, cybersecurity

References

1. Mouna J., Latifa B., Latifa B.R., Anis A. Classification of security threats in information systems. Procedia Computer science, 2014, vol. 32, p. 489-496.
2. Derendyaev D.A., Gatchin Yu.A., Bezrukov V.A. Opredelenie vliyaniya chelovecheskogo faktora na osnovnye kharakteristiki ugroz bezopasnosti [Determining the influence of the human factor on the main characteristics of security threats]. Kibernetika i programmirovaniye [Cybernetics and programming], 2019, no. 3, pp. 38-42.

3. Gyorodi R., Cornelia G., Pecherle G., Radu L. Network security using firewalls. Journal of Computer science and control systems, 2008, vol. 1.
4. Kao D.Y., Wang S.J., Huang F. Dataset Analysis of proxy logs detecting to curb propagations in network attacks. Intelligence and security informatics, 2008, pp. 245-250.
5. Marshall B., Chen H. Using importance flooding to identify interesting networks of criminal activity. Lecture notes in computer science, 2006, vol. 3975, pp. 14-25.
6. Mukkamala S., Sung A. Identifying significant features for network forensic analysis using artificial techniques. International journal of Digital evidence, 2003, vol. 1, no 4.
7. Bayraktar C., Karakaya Z., Gökçen H. Real-time anomaly detection system within the scope of smart factories. The journal of Supercomputing, 2023, vol. 79.
8. Wang C., Zhou H., Hao Z., Hu S., Li J., Zhang X., Jiang B., Chen X., Network traffic analysis over clustering-based collective anomaly detection. Computer networks, 2022, vol. 205, pp. 108760.
9. Yang T., Jiang Z., Liu P., Yang Q., Wang W. A traffic anomaly detection approach based on unsupervised learning for industrial cyber-physical system. Knowledge-Based Systems, 2023, vol. 279, pp. 110949.
10. Demertzis K., Tsiknas K., Taketzis D., Skianis C., Iliadis L. Darknet Traffic big-data analysis and network management for real-time automating of the malicious intent detection process by a weight agnostic neural networks framework. Electronics, 2021, vol. 10, pp. 781.
11. Isaev S.V. Kiberbezopasnost' nauchnogo uchrezhdeniya - aktivy i ugrozy [Cybersecurity of a scientific institution - assets and threats]. Informatizatsiya i svyaz' [Informatization and communication], 2015, vol 1, p. 53-57.
12. Isaev S.V. Analiz kiberugroz i ikh istochnikov dlya korporativnoy seti Krasnoyarskogo nauchnogo tsentra SO RAN [Analysis of cyber threats and their sources on the corporate network Krasnoyarsk Scientific Center of the SB RAS]. Informatsionnyie i matematicheskie tehnologii v nauke i upravlenii [Information and mathematical technologies in science and management], 2016. № 4-1. p. 76-85. (In Russ.)
13. Blei D.M. Probabilistic topic models. Communications of the ACM, 2012, vol. 55, no. 4, pp. 77-84.
14. Fei, B., Eloff, J., Oliver, M., Venter, H. Analysis of web proxy logs. IFIP international conference on digital forensics, Orlando, 2006, vol. 222, pp. 247-258.
15. Scott D.W. Multivariate density estimation. Theory. Practice and visualization: second edition. New York, 2015.
16. King T.L., Bentley R.J., Thornton L.E. [et al.] Using kernel density estimation to understand the influence of neighbourhood destinations on BMI. BMJ Open, 2016, vol. 6.
17. Kalinic M., Krisp J. Kernel density estimation (KDE) vs. hot-spot analysis - Detecting criminal hot spots in the city of San Francisco. Lund, Sweden, 2018.
18. Vorontsov K. V. Obzor veroyatnostnykh tematicheskikh modelei [Overview of probabilistic thematic models], 2021
19. Albalawi R., Yeap T., Benyoucef M. Using topic modeling methods for short-text data: A comparative analysis. / Frontiers in artificial intelligence, 2020, vol. 3.
20. Jelodar H., Wang Y., Yuan, Ch., Xia, F. Latent dirichlet allocation (LDA) and topic modeling: models, applications, a survey, 2017.
21. Tharwat A., Gaber T., Ibrahim A., Hassanien A.E. Linear discriminant analysis: A detailed tutorial. Ai communications, 2017, vol. 30, pp. 169-190.

Isaev Sergey Vladislavovich. Cand. Sc., docent, institute of computational modeling SB RAS, deputy director for research, head of the department of information and telecommunication technologies. Research interests: cybersecurity and information protection, Internet technologies, distributed information systems, intelligent systems. AuthorID: 1623, SPIN: 1552-8542, si@icm.krasn.ru, 660036, Russia, Krasnoyarsk, Akademgorodok, bld. 50, p. 44, ICM SB RAS.

Dontsov Denis Yurievich. Postgraduate student, institute of computational modeling SB RAS. Research interests: machine learning, Internet technologies. AuthorID: 1623, doncov.dy@icm.krasn.ru, 660036, Russia, Krasnoyarsk, Akademgorodok, bld. 50, p. 44, ICM SB RAS.

Статья поступила в редакцию 05.05.2023; одобрена после рецензирования 07.12.2023; принята к публикации 07.12.2023.

The article was submitted 05/05/2023; approved after reviewing 12/07/2023; accepted for publication 12/07/2023.

