

ISSN 2413 - 0133
Scientific journal

ИНФОРМАЦИОННЫЕ И МАТЕМАТИЧЕСКИЕ ТЕХНОЛОГИИ В НАУКЕ И УПРАВЛЕНИИ

№3(35)/2024

ИНФОРМАЦИОННЫЕ И МАТЕМАТИЧЕСКИЕ ТЕХНОЛОГИИ В НАУКЕ И УПРАВЛЕНИИ

Научный журнал
№ 3 (35)



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.
 Marchenko M.A.
 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, Таганрог, ЮФУ
 Лис Р., Польша, Wrocław University of Science and Technology
 Марченко М.А. д.ф.-м.н., профессор РАН, Новосибирск, ИВМиМГ СО РАН
 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.
 Tech. 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»
 Адрес типографии: 664046, г. Иркутск, ул. Волжская, 14
 *Подписка на журнал доступна со второго полугодия 2021 г.
 Дата выхода 17.10.2024 г. Тираж 50 экз.

© Издательство ИСЭМ СО РАН
 Цена свободная (12+)

Методологические аспекты информационных и математических технологий

- Атаева О.М., Массель Л.В., Серебряков В.А., Тучкова Н.П.**, Интеллектуальный анализ данных при построении графа знаний мультидисциплинарного журнала 5
- Сидорова Е.А., Кононенко И.С.** Онтологический анализ приемов аргументации в научном дискурсе 20

Методы, технологии и приложения искусственного интеллекта

- Дорогов А.Ю.** Пластичность самоподобных нейронных сетей 33
- Милантьев С.А., Святкина В.И., Бессмертный И.А., Зайченко К.В.** Анализ снимков кожных новообразований с применением комбинированной архитектуры сверточных нейронных сетей 44
- Подобедова Е.В., Жукова М.С., Аршинский Л.В.** Экспертная система поддержки расследования происшествий на железнодорожном транспорте 57

Математические, информационные и интеллектуальные технологии в энергетике

- Баденко В.В., Козлов А.Н.** Первичный анализ работы гибридной микросети лабораторного масштаба с использованием системы мониторинга 73
- Елисеев А.В., Миронов А.С.** Элементы концепции структурно-функционального подхода к проблеме управления динамическими состояниями технических объектов в условиях вибрационных нагрузений 82
- Дунаев М.П., Саушкин А.А.** Моделирование трехтактного частотно-импульсного преобразователя 94

Цифровая экономика и управление

- Воскобойников М.Л., Феоктистов А.Г.** Сравнительный анализ систем управления научными рабочими процессами 102
- Краковский Ю.М., Киргизбаев В.П.** Влияние вероятностных моделей работ, связанных с защитой информации, на значения показателей эффективности 112
- Камалетдинова Л.Р., Романов А.А., Филиппов А.А., Ярушкина Н.Г.** Формирование базы правил для PLM систем 120

Математическое моделирование и его применение в научных исследованиях

- Дудаев М.А.** Модификация алгоритма минимальной степени для повышения эффективности вычислений при итерационном решении контактной задачи теории упругости методом конечных элементов 133
- Маланова Т.В., Бахвалов С.В., Янхаев Д.В., Аршинский В.Л.** Планирование поставок продукции на год с учетом минимизации затрат на перевозку для бесперебойного снабжения потребителей 148
- Колокольцева И.М., Иваньо Я.М.** Параметрическая оптимизация производства аграрной продукции при влиянии на урожай саранчовых вредителей 157
- Шемахин А.Ю., Абдуллин И.Ш., Желтухин В.С., Макеева Е.А.** Характеристики струйного высокочастотного индукционного разряда пониженного давления в процессах обработки твердых тел 167

- Приложение.** Основные рубрики журнала «Информационные и математические технологии в науке и управлении» и их соответствие специальностям ВАК 176

Methodological aspects of information and mathematical technologies

Ataeva O.M., Massel L.V., Serebryakov V.A., Tuchkova N.P. Data mining when constructing a knowledge graph of a multidisciplinary journal 5

Sidorova E.A., Kononenko I.S. Ontological analysis of argumentation techniques in scientific discourse 20

Methods, technologies and applications of artificial intelligence

Dorogov A.Yu. Plasticity of self-similar neural networks 33

Milantev S.A., Sviatkina V.I., Bessmertny I.A., Zaichenko K.V. Image analysis of skin lesion using a combined convolutional neural network architecture 44

Podobedova E.V., Zhukova M.S., Arshinskiy L.V. Expert system for the support of investigation of accidents in railway transport 57

Mathematical, information and intellectual technologies in the energy sector

Badenko V.V., Kozlov A.N. Initial analysis of the performance of a laboratory-scale hybrid microgrid using a monitoring system 73

Eliseev A.V., Mironov A.S. Elements of the concept of a structural and functional approach to the problem of controlling the dynamic states of technical objects under vibration loads 82

Dunaev M.P., Saushkin A.A. Modeling of a three-stroke frequency-pulse converter 94

Digital economy and management

Voskoboinikov M.L., Feoktistov A.G. Comparative analysis of scientific workflow management systems 102

Krakovsky Y.M., Kirgizbaev V.P. The impact of probabilistic models of works related to information protection on the values of performance indicators 112

Kamaletdinova L.R., Romanov A.A., Filippov A.A., Yarushkina N.G. Rule base forming for PLM systems 120

Mathematical modeling and its application in scientific research

Dudaev M.A. The minimum degree algorithm modification for improving the computational efficiency in the iterative solution of the contact problem of elasticity theory using the finite element method 133

Malanova T.V., Bakhvalov S.V., Yanhaev D.V., Arshinsky V.L. Planning product supplies for the year, taking into account minimizing transportation costs 148

Ivanyo Ya.M., Kolokoltseva I.M. Parametric optimization of agricultural production under the influence of locust pests on the harvest 157

Shemakhin A.Yu., Abdullin I.Sh., Zheltukhin V.S., Makeeva E.A. Characteristics of a intermediate pressure radio frequency inductive coupled stream discharge for solid processing 167

Appendix. Main headings of the journal "Information and Mathematical Technologies in Science and Management" and their correspondence to the specialties of the Higher Attestation Commission of the Russian Federation 176

Методологические аспекты информационных и математических технологий

УДК 004.8+519.217 + 519.876.2

DOI:10.25729/ESI.2024.35.3.001

Интеллектуальный анализ данных при построении графа знаний мультидисциплинарного журнала

Атаева Ольга Муратовна¹, Массель Людмила Васильевна², Серебряков Владимир Алексеевич¹, Тучкова Наталия Павловна¹

¹ФИЦ «Информатика и управление» РАН, Россия, Москва, oli.ataeva@gmail.com

²Институт систем энергетики им. Л.А. Мелентьева СО РАН, Россия, Иркутск

Аннотация. В работе исследуется тематическое многообразие междисциплинарного журнала. Цель исследований составляет построение графа знаний журнала для тематического представления и систематизации электронного архива и новых публикаций журнала. Исходные данные представляют собой статьи журнала, посвященные различным информационным и математическим технологиям в науке и управлении, то есть междисциплинарным исследованиям. Предлагается систематизация текстов с помощью методов векторного анализа. В процессе тематического анализа контента журнала предлагается разбиение на рубрики, устанавливаются связи рубрик и статей с соответствующими описаниями специальностей ВАК. Для анализа тематики используется разведочный анализ исходных текстов, далее применяются методы интеллектуального анализа данных. Результаты разбиения предоставляются экспертам журнала, после чего вырабатывается решение о формировании тематической рубрики и включении в нее специальностей ВАК. Статьи журнала интегрируются в семантическую библиотеку LibMeta, в силу чего онтология библиотеки достраивается и формируется онтология журнала, и на этой основе строится граф знаний журнала. Предлагается процедура навигации по контенту журнала с помощью графа знаний в семантической библиотеке LibMeta, которая может стать основой для информационного сопровождения научных исследований и создания цифрового ассистента в междисциплинарной предметной области. Примеры приведены для конкретного контента журнала, но предложенная технология может быть распространена на другие журналы, так как большинство журналов, относящихся к нескольким специальностям ВАК, естественным образом захватывают несколько дисциплин.

Ключевые слова: граф знаний, семантическая библиотека, достраивание онтологии, кластеризация научных статей, суммаризация текста

Цитирование: Атаева О.М. Интеллектуальный анализ данных при построении графа знаний мультидисциплинарного журнала / О.М. Атаева, Л.В. Массель, В.А. Серебряков, Н.П. Тучкова // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 5-19. – DOI: 10.25729/ESI.2024.35.3.001.

Введение. Задача тематической кластеризации обсуждается достаточно давно, начиная с создания первых библиографических массивов данных [1]. Для анализа научных публикаций эта задача будет ставиться снова и снова в процессе развития науки и появления новых предметов исследования. Кластеризация данных предполагает группировку данных на основе признаков им сходств без заранее определенных категорий [2, 3]. Наиболее известные научные библиографические ресурсы, такие как, WoS (<https://clarivate.com/cis/solutions/web-of-science/>), Scopus (<https://www.scopus.com/>), eLIBRARY (<https://www.elibrary.ru/>), предлагают поиск по ключевым словам и другим метаданным, которые, в свою очередь, связаны как с публикацией, так и с конкретной предметной областью (Про). В целом, эти ресурсы не предлагают сервиса по анализу содержания публикаций, исключение составляет zbMATH Open (<https://zbmath.org/>), где можно найти отзывы и рецензии. В то же время нарастающий объем публикаций оставляет необходимость в предварительной оценке тематики публикации прежде, чем купить или скачать ее для ознакомления. Работы по кластеризации научных тек-

стов с целью их тематического распределения в семантической среде продолжают и нарастают с применением средств машинного обучения (МО) и моделей естественного языка [4, 5, 6]. Определенный скачок в этом направлении произошел в связи с применением векторных алгоритмов [7]. Почти через полвека появилась технология построения семантических векторных представлений, или эмбедингов (embedding) [8, 9, 10], в том числе и для научных областей. Так, для eLIBRARY разработана нейронная сеть для получения эмбедингов научных текстов [11, 12, 13] на основе аннотаций к статьям, с помощью которых можно решать различные задачи обработки текстов. Тем не менее, далеко не все ПрО и не все журналы в полном объеме охвачены этой технологией и проблема тематического поиска остается, а ее актуальность неизбежно связана с экспоненциальным ростом числа публикаций. К тому же есть междисциплинарные ПрО, для которых тоже необходимо составить картину принадлежности к множеству ПрО. Предлагаемое исследование посвящено именно междисциплинарной тематике журнала и кластеризации этого контента с целью выявления множеств научных статей, которые можно отнести к разным ПрО.

Новый этап развития методов искусственного интеллекта (ИИ) связан с появлением новых эффективных нейросетевых методов для обработки естественного языка. Но, несмотря на это, особые трудности возникают при обработке научных статей из-за специфичной для ПрО терминологии и сложных идей, часто встречающихся в научной литературе [14,15]. Однако, использование графов знаний (ГЗ), основанных на онтологии ПрО, в совокупности с размеченными статьями при тонкой настройке нейросетевых методов, позволяет улучшить результаты обработки и извлечения смысла текстов, а поиск становится более обоснованным [16, 17, 18].

ГЗ обычно представляется как набор объектов, связанных между собой. Объекты представляют собой узлы этого графа, а связи представляются в виде ребер. Для некоторой ПрО ГЗ ограничивается определенным типом объектов и заданным набором связей. Формально при этом ПрО задается в виде онтологии $\langle R, A, C, I \rangle$, которая представляет собой типы объектов R , их атрибуты A и отношения C , а также функции интерпретации I этих отношений. Таким образом, множество $\{R, A, C, I\}$ задает описание структуры ПрО.

Представление знаний в виде онтологии ПрО и навигация с помощью ГЗ представляет собой технологию поиска, при которой можно оставаться в рамках тематики и опираться на достоверные данные, связанные с энциклопедиями, словарями, тезаурусами, классификаторами и первоисточниками.

Работа направлена именно на задачу построения ГЗ *междисциплинарной* ПрО контента журнала (лов) [17, 18]. С этой целью был проведен интеллектуальный анализ поступающих публикаций для определения их тематической принадлежности. Интеллектуальный анализ текстов подразумевает совместную работу специалистов ПрО и средств ИИ. Для описания ПрО привлекается контент семантической библиотеки LibMeta (libmeta.ru) и классические первоисточники, рекомендованные экспертами. Начальное описание ПрО, как семантически связанной структуры данных, опирается на классификаторы УДК, MSC, рубрики журналов, специальности ВАК журналов, специальные словари и предметные указатели монографий.

Построенный ГЗ журнала интегрируется с энциклопедиями LibMeta и онтологией ИИ в LibMeta. Добавление контента журнала влечет за собой процедуру доработки онтологии семантической библиотеки.

1. Граф знаний журнала. В работе исследуются проблемы построения и наполнения ГЗ на примере семантической библиотеки LibMeta. Используется подход онтологического проектирования для построения ГЗ. При построении онтологии графа определяются типы узлов и связи. При наполнении графа предполагается добавление новых узлов и связей в соответствии с описанием в онтологии и в соответствии с добавленным контентом журнала.

Научный журнал может быть посвящен отдельной теме (научной дисциплине) или охватывать ряд предметов (междисциплинарность). И в том, и другом случае тематика журнала, как правило, подпадает под несколько специальностей ВАК. В силу развития науки в целом выбор специальностей зависит от тенденций в исследованиях [19, 20] и от позиции редакции. Научная статья – это исследование, которое посвящено какой-то проблеме (задаче) и предполагает структуру изложения, из которой читателю должна быть понятна суть проблемы и пути ее исследования. Современные библиографические базы данных предлагают различные варианты поиска публикаций, включая использование отзывов, ссылок на аналогичные (с позиции базы данных) публикации и пр. В целом, это путь для помощи пользователю в навигации по библиографическим данным. Большие информационные поисковые ресурсы (google и др.) используют связи данных и строят для навигации ГЗ. Эта технология, основанная на семантическом анализе контента и нейросетевых алгоритмах, также применима для полнотекстовых научных библиотек, а именно, информационного сопровождения автора, создания цифрового ассистента научных исследований.

Цифровой ассистент (семантическая библиотека) выполняет задачу информационного сопровождения автора и редактора в подборе УДК, ключевых слов, аналогичных статей, суммаризации, автоматического аннотирования и др.

Для реализации подхода навигации по научному контенту с помощью ГЗ требуется значительная подготовка текстов, что составляет существенное отличие от текстов, с которыми работают поисковые агрегаторы в интернет. Несмотря на известную структуру научных статей, в них необходимо выделить *предмет* исследования и *идеи* его реализации [14]. Поэтому особую важность приобретает тематическая кластеризация и соответствие специальностям Высшей аттестационной комиссии (ВАК).

Логика построения научных публикаций журнала, как правило, организована таким образом, что в тексте присутствуют выражения типа: «в работе исследуется **задача**», «работа посвящена **проблеме**», или близкие варианты. Далее следуют **методы** исследования, **результаты**, **примеры** и т.д. В контексте специальностей ВАК ПрО журнала включает описание множества задач, соответствующих выбранным специальностям. С другой стороны, тематические рубрики журнала также отражают позицию редакции к соответствию специальностям конкретных задач, отраженных в статьях. Таким образом, ПрО журнала должна включать описание множества задач, соответствующих, с точки зрения редакции, как специальностям ВАК, так и предметным рубрикам журнала. При таком взгляде на ПрО журнала вершинами ГЗ журнала могут рассматриваться тематические рубрики. Рубрики могут быть связаны с несколькими специальностями ВАК и включать описания задач, отраженных в статьях, отнесенных редакцией к этим рубрикам. Перемещение по таким узлам графа будет означать навигацию по ПрО журнала, в том числе по рубрикам и специальностям ВАК, относящимся к рубрикам. Эта навигация может быть применена как для журналов, посвященных одной научной дисциплине, так и для междисциплинарных журналов, и отличается от навигации по библиографическим данным, которая характерна для большинства научных библиографических баз данных.

Тематическая кластеризация в исследовании реализована, как **распределение статей по рубрикам и распределение по специальностям ВАК**. Эти две задачи включают этапы подготовки данных, обработку данных с помощью методов МО и дальнейший анализ данных для их интеграции в цифровую библиотеку путем доработки онтологии библиотеки и построения на ее основе ГЗ ПрО журнала. На рисунке 1 представлен пример ГЗ ПрО журнала после интеграции и достраивания онтологии LibMeta. На первом уровне представлены общие понятия системного тезауруса, на втором описываются понятия ПрО, которые определяют типы узлов

в ГЗ ПрО, на третьем уровне задаются отношения между ними. На четвертом уровне — представлен уровень данных или ГЗ. На рисунке 1 приведена информация в ГЗ о публикации и показаны некоторые ее связи с типами узлов и конкретные данные (специальность, рубрика журнала, название статьи и т.д.).

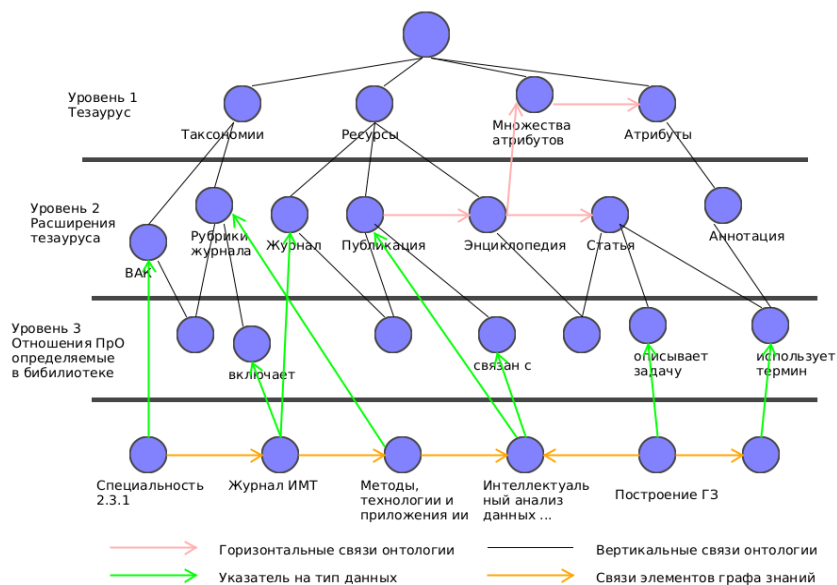


Рис. 1. Пример ГЗ журнала в семантической библиотеке LibMeta

2. Распределение по рубрикам и тематическим кластерам.

2.1. Разведочный анализ данных журнала. Предварительный анализ контента журнала и начальных рубрик выявил проблемы дублирования и широкий тематический охват исследований, опубликованных в статьях.

Для выполнения качественной кластеризации данных необходимо привлечение эксперта ПрО для оценки результата автоматической обработки с помощью алгоритмов *векторизации* [7, 8, 9], *LDA* (Latent Dirichlet Allocation [21]) и *k-means* [22]. Выявление рубрик и распределение публикаций по новым рубрикам позволит далее связать их явно со специальностями ВАК. Надо отметить, что особенность междисциплинарных журналов заключается именно в наличии статей из *различных научных дисциплин*, объединенных одной общей идеей, декламируемой издателями, которая позволяет публиковать их совместно в одном журнале. Это приводит к тому, что в журнале может обнаружиться очень *много тематических рубрик*, в том числе как достаточно близких, так и различных. В то же время все *статьи должны подчиняться специальностям ВАК*, которые в свою очередь также *допускают довольно широкий тематический диапазон*. Однако, читателю важно понимать какому разделу (рубрике) принадлежит публикация, так как с этим связан выбор статьи для ознакомления, и, в конечном счете, скорость и удобство этого выбора. Значит, в журнале должно быть некоторое разумное, *оптимальное* с точки зрения издателей и читателей количество рубрик, которое отражает, в то же время, тематику статей.

Для определения *оптимального* количества тематических рубрик необходимо подготовить исходные массивы данных и выбрать параметры их автоматической обработки с помощью алгоритмов ИИ и провести разведочный анализ данных журнала.

2.1.1. Подготовка данных. Данные публикаций изначально были представлены в слабоструктурированном виде, в процессе предобработки из них были выделены основные метаданные, такие, как автор, название, аннотация и заголовок и т.д.

Предобработка текстов (очистка) включала работу по токенизации, лемматизации (данные лемматизируются, биграмы добавляются, убираются часто встречающиеся слова, и

слова, которые встречаются меньше 5 раз, убираются имена собственные и т.д.), удалении стоп-слов и разбиении текста на разделы.

Далее с помощью статистических векторных методов, основанных на подсчете *TF-IDF* (Term Frequency-Inverse Document Frequency [23]), были выделены ключевые слова, в частности, ключевые слова, которые были предложены для индексации отдельных документов, где авторские ключевые слова не распознались или отсутствовали. Ниже приведен пример (Пример 1) извлечения из статьи ключевых слов *при предобработке* и *авторских* ключевых слов.

Пример 1.

Тучкова Н. П., 'Атаева О. М. Подходы к извлечению знаний в научных предметных областях
Ключевые слова: ['тезаурус', 'предметный', 'предметный область', 'знание', 'извлечение знаний', 'область', 'извлечение', 'публикация', 'данные', 'интеллект', 'искусственный', 'метрика', 'искусственный интеллект', 'наукометрический', 'онтология тезаурус', 'цифровой', 'структурирование', 'научный', 'сохранение извлечение', 'наукометрический показатель']

Авторские ключевые слова из статьи: ['структурирование данных', 'тезаурус предметной области', 'метрики', 'тезаурус оду']

2.1.2. Распределение по рубрикам. Начальное разбиение контента журнала за весь период публикаций состояло из 53 рубрик, которые содержали тематические пересечения и дублирование. После предварительной обработки и очистки был произведен анализ текстов и их разбиение на кластеры.

В процессе исследования данных было построено *векторное представление* документов (контента журнала) с помощью статистической меры *TF-IDF* и на основе этого представления было проведено разбиение на кластеры с помощью алгоритма *LDA*. Анализ результатов подтвердил предположение о том, что изначальная рубрикация была избыточна. Было выявлено, что документы часто не попадали в одну рубрику, а распределялись по 5 и более рубрикам.

2.1.3. Оценка разбиения на кластеры. Следующий шаг в исследованиях был направлен на определение примерного «оптимального» количества тематических кластеров. Была применена метрика согласованности тем (Topic coherence) как показатель связности при тематическом моделировании для того чтобы оценить, насколько интерпретируемы темы. В этом случае *темы* представлены как первые *N* слов с наибольшей вероятностью принадлежащих к некой конкретной теме. То есть, показатель связности свидетельствует от том, насколько эти слова *похожи друг на друга* для разного количества тем.

В диапазоне от 5 до 7 расстояния между кластерами перестают сильно изменяться, как видно на рисунке 2.

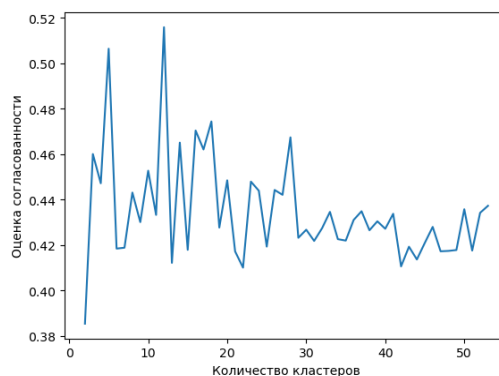


Рис. 2. Кривая оценки согласованности тематик

2.1.3. Метод k-means. Следующий эксперимент был проведен с использованием алгоритма *k-means* с использованием эмбедингов на основе нейросетевой модели *SciRus-tiny* [19], обученной на научных текстах. Модель имеет небольшое количество параметров и невысокие требования к вычислительным ресурсам.

Для нахождения оптимального количества кластеров была использована функция оценки расстояния между кластерами на основе косинусного расстояния (рисунок 3). Эксперимент показывает, что использование эмбедингов научных текстов на нейросетевых моделях на данных журнала не сильно сказалось на результате. Для интерпретации полученных результатов каждый кластер более детально рассмотрен экспертами. Оптимальное количество кластеров было определено равным 12-и.

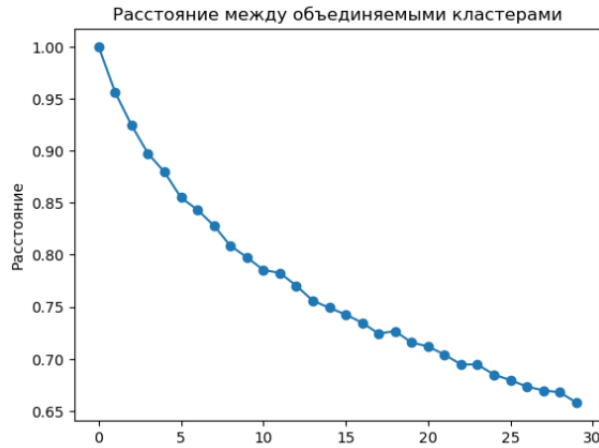


Рис. 3. Оценка расстояния между кластерами на основе косинусного расстояния

Гиперпараметры моделей подбирались на основе анализа и исследования степени соответствия построенной кластерной структуры тематической модели, при этом использована метрика *Silhouette Coefficient*.

На рисунке 4 значение кластеров 5 и 10, выглядят *лучше*, в том смысле, что оценка силуэта кластера больше среднего значения для каждого кластера. Для значений 15 и 20 разбиение *не оптимально*, так как присутствуют кластеры с оценкой меньше среднего, есть большие колебания в размере графика. На рисунке 5 приведена визуализация разбиения на кластеры для $k = 5, 10, 15$.

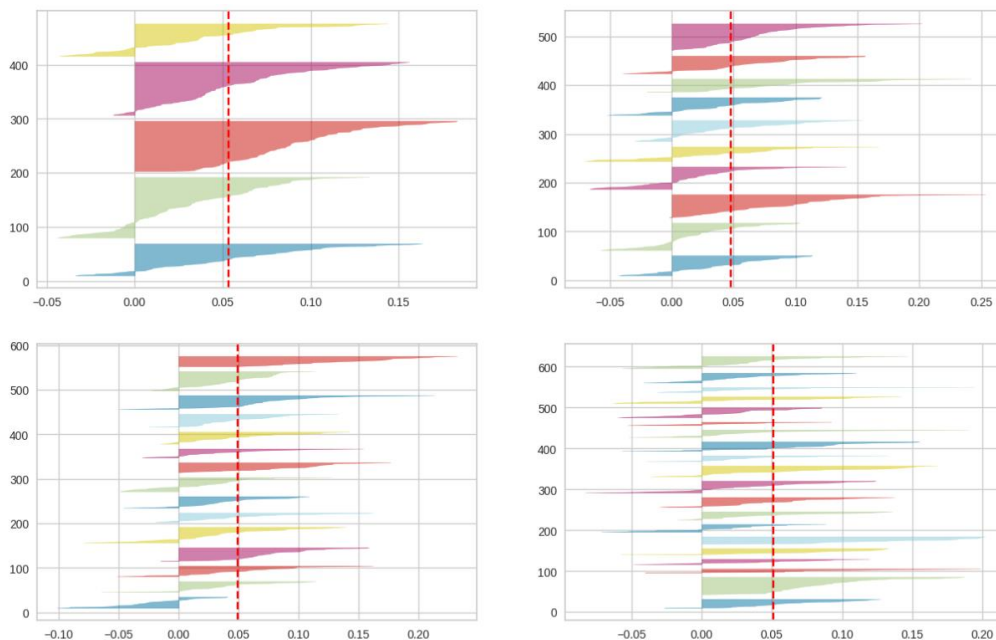


Рис. 4. Оценка расстояния между кластерами на основе косинусного расстояния

Полученные результаты были предоставлены для анализа экспертам и редакционной коллегии журнала для подтверждения и коррекции. Каждый тематический кластер представляется в виде следующих множеств данных: набора ключевых слов и словосочетаний, которые

определяются с помощью статистических методов, а также отдельного набора авторских ключевых слов из статей, которые были отнесены к этому кластеру и списка самих статей.

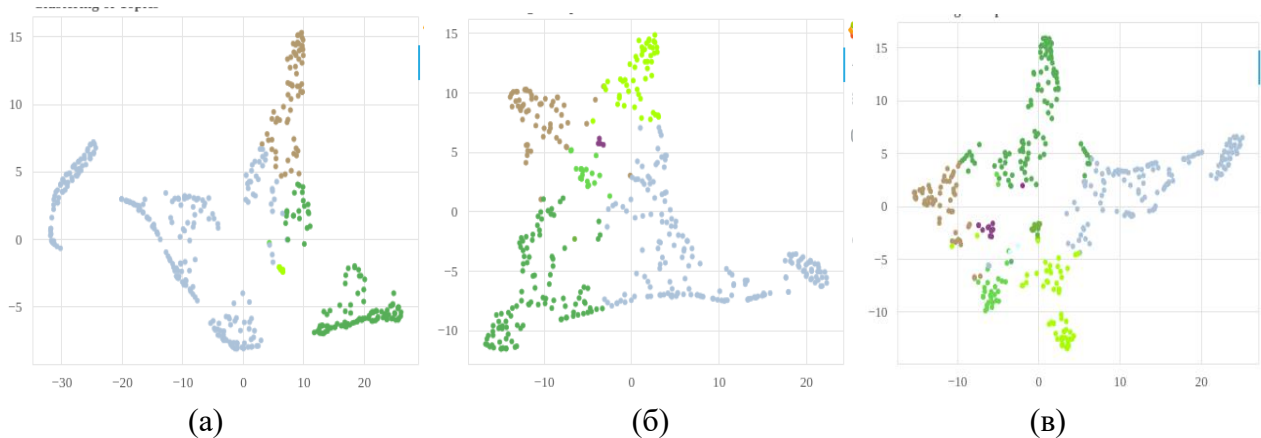


Рис. 5. Распределение статей по темам (*k-means*): а) $k = 5$, б) $k = 10$, в) $k = 15$

2.2. Достаивание онтологии. На основе проведенной работы и оценки экспертов были выделены 7 тематических рубрик, которые добавлены в онтологию и связаны с классификаторами (ВАК, УДК), тезаурусом, энциклопедиями, публикациями, ключевыми словами, что иллюстрируется на рисунке 6.

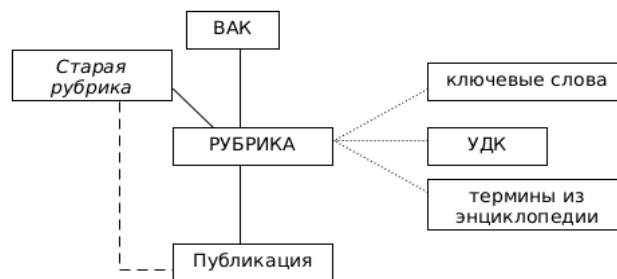


Рис. 6. Схема связей публикации с рубриками журнала и метаданными

Пунктирной линией обозначены связи предыдущего разбиения публикаций на начальные «старые» рубрики. После создания новых рубрик, каждая из них благодаря связям публикаций со «старыми» рубриками, была наполнена ключевыми словами, связями УДК, терминами энциклопедии.

3. Распределение по специальностям ВАК.

3.1. Подготовка данных. Для этого этапа необходимо проанализировать средствами *NLP* [24, 25] не только контент журнала, но и содержание паспорта ВАК. В разделах, которые содержат описание специальности ВАК, прослеживается четкая структура предложений, где выделяются области (сферы) исследования, цели, предмета исследования, средства исследования. Это позволяет проанализировать содержание паспорта ВАК с помощью лингвистических правил.

3.2. Методы. Разбор предложений на структуры был выполнен с помощью лингвистических правил [20], где было выделено: «что делается», «какими методами», «в какой области» и «какой результат должен получиться». Для этого была проанализирована структура предложения, определены связанные слова с применением частеречных меток (*POS-тэги*рование, *part-of-speech tagging*), и т.д. Схематически этот процесс разбора содержания паспорта специальности иллюстрируется на рисунке 7.

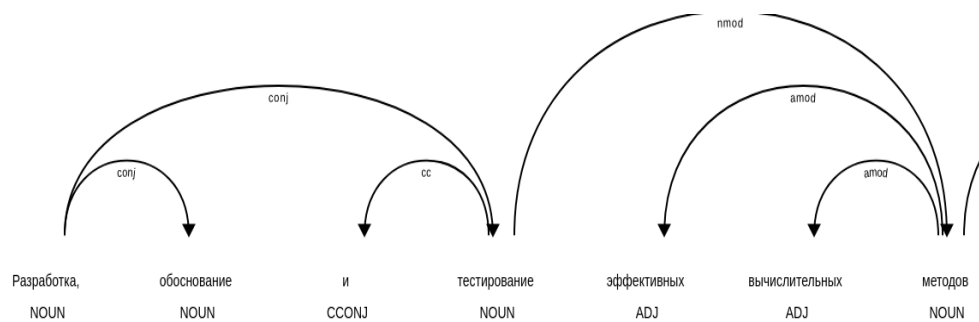


Рис. 7. Связи в предложениях описания специальностей ВАК на основе лингвистических правил

3.3. Анализ результатов. Разбор предложений на структуры позволил увидеть и сформировать для описания специальности ВАК словари общей лексики, которые включают в себя, например, описание для всех специальностей: задачи, методы, решения.

Для конкретных специальностей эта общая лексика уточняется, например, для специальности 2.3.1 рассматриваются решения и методы для задач оптимизации, управления, принятия решений, обработки информации, искусственного интеллекта.

На рисунке 8 приведено описание паспорта ВАК специальности 1.2.2. В каждом предложении соответствующими цветами выделены: область (серый), задача (желтый), предмет (розовый) и инструмент исследования (фиолетовый).

Паспорт научной специальности 1.2.2. «Математическое моделирование, численные методы и комплексы программ»	
1.	Разработка новых математических методов моделирования объектов и явлений
2.	Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий.
3.	Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента
4.	Разработка новых математических методов и алгоритмов интерпретации натурального эксперимента на основе его математической модели.
5.	Разработка новых математических методов и алгоритмов валидации математических моделей объектов на основе данных натурального эксперимента или на основе анализа математических моделей.

Рис. 8. Пример анализа паспорта ВАК для специальности 1.2.2.

3.4. Достаивание онтологии. В онтологию LibMeta добавлен тезаурус для поддержки описания специальностей ВАК, модель которого включает соответствующие связи и атрибуты, и загружена информация о специальностях, соответствующих контенту журнала (рисунок 9).

На рисунке 9 представлен фрагмент информационной модели специальности ВАК, которая включает свойства «Область исследования», «Предмет исследования», «Инструмент исследования». Анализ паспортов ВАК позволил выделить эти свойства, как общие для всех рассматриваемых специальностей. Некоторые конкретные значения этих свойств из специальности 1.2.2 приведены на рисунке 9.

В результате анализа публикации и определения ее специальности для каждой публикации получаем соответствующее описание по разделам специальности ВАК. Более того, формальные описания специальностей наполняются в тезаурусе LibMeta конкретными ключевыми словами из публикаций журнала. Это позволяет автоматически связывать новые публикации со специальностями для дальнейшей оценки экспертами, рецензентами и авторами.

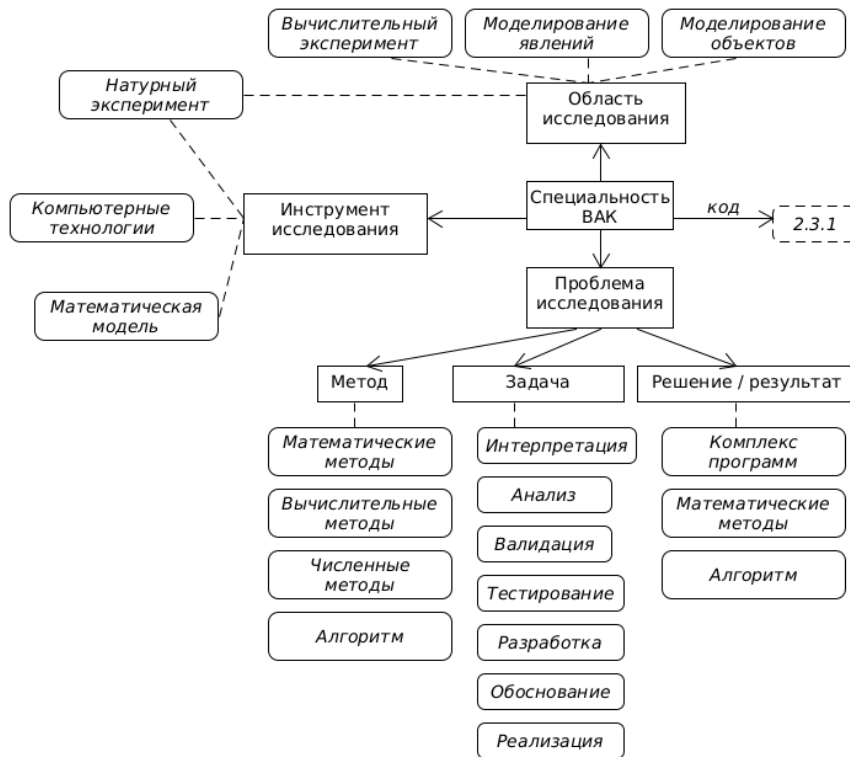


Рис. 9. Схема специальности ВАК 1.2.2 для журнала

4. Методология. Исходные данные представляют собой статьи журнала, посвященные различным информационным и математическим технологиям в науке и управлении, то есть междисциплинарным исследованиям. Большинство журналов, относящихся к нескольким специальностям ВАК, естественным образом захватывают несколько дисциплин. В работе предложен подход к тематической кластеризации на примере одного журнала¹, но используются методы, не ограничивающие его общность, что позволяет говорить об общей методологии в задаче тематической кластеризации таких массивов данных. Результаты предложенных исследований носят прикладной характер, поскольку позволяют организовать навигацию по контенту журнала и использованием графа знаний.

Основная идея, которая была заложена в предлагаемом подходе, – это использовать технологии, которые позволяют работать с «сырым» материалом русскоязычных научных публикаций, выделяя из текстов «задачи», «методы», «решения», «результаты» и помещать их в семантическую библиотеку. Для этого выделяются тематические кластеры и материал разбивается по рубрикам, узлам графа. Множеству рубрик соответствуют ключевые слова, входящие в специальности ВАК. Специальности ВАК – это тоже узлы графа. Специальности ВАК, кроме собственных терминов, приобретают множество ключевых слов из статей, входящих в рубрики. Таким образом, статьи необходимо подготовить, а именно, *выделить задачи, методы их исследования и результаты* с помощью средств обработки естественного языка и консультации экспертов. Такая разметка возможна только совместно с экспертами, если не присутствует в самом тексте статей.

Про журнала, таким образом, представлена терминологией, включая описание задач, которым посвящены публикации, при этом вершинами графа Про будут *рубрики* журнала и *специальности* ВАК. Перемещаясь по вершинам графа, получаем множества задач, методов,

¹ Рассматривался журнал «Информационные и математические технологии в науке и управлении». В Приложении (после статей этого выпуска) приведен список принятых, после выполненного анализа, рубрик журнала и соответствующих им специальностей ВАК

результатов и т.д., которые относятся к статьям в рубрике, и далее, публикации, в которых изучается та или иная задача междисциплинарной ПрО. Полученный информационный ресурс отличается от того, как журнал присутствует в интернете на основе библиографических данных. Результат (рисунки 10, 11) позволяет быстро *посмотреть, нужна ли пользователю конкретная статья или нет, не скачивая, не читая целиком*. Применение результатов предполагает создание рекомендательной системы для информационного сопровождения пользователя, автора, редактора, рецензента и др.

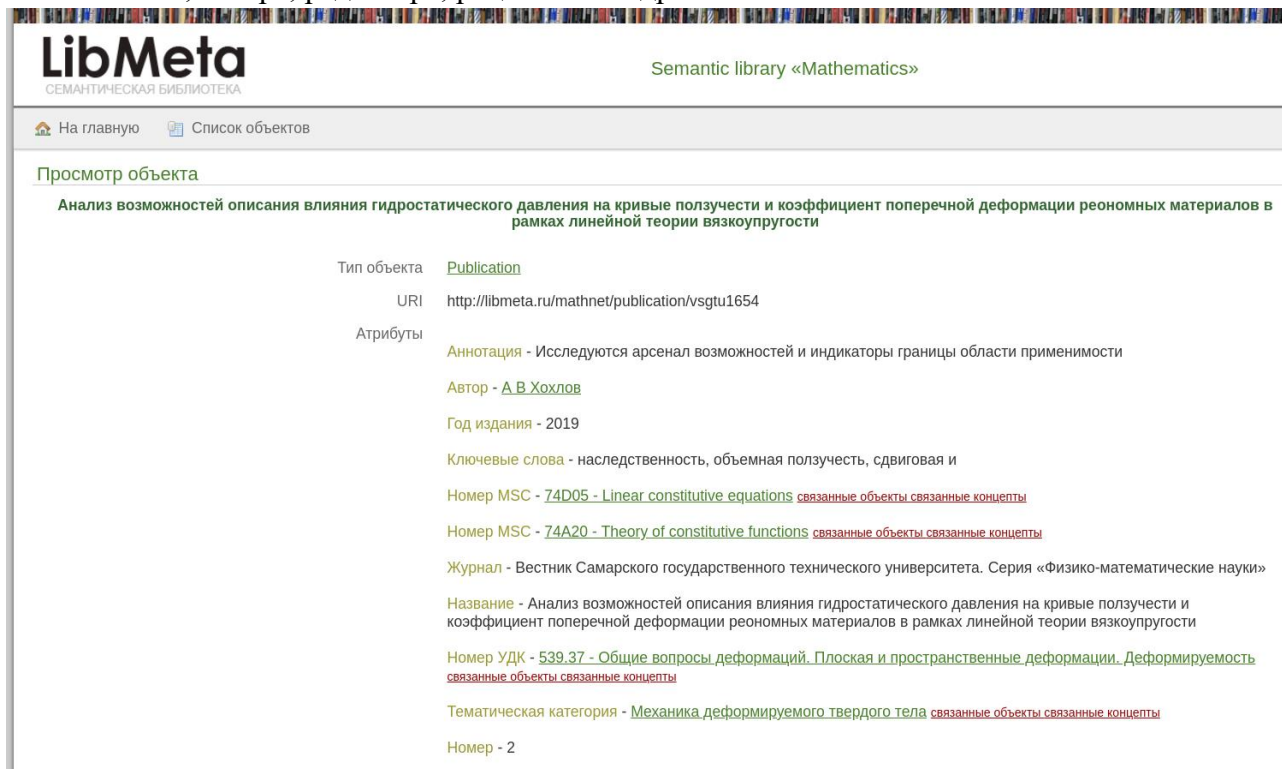


Рис. 10. Скрин 1. Публикация и тематические разделы

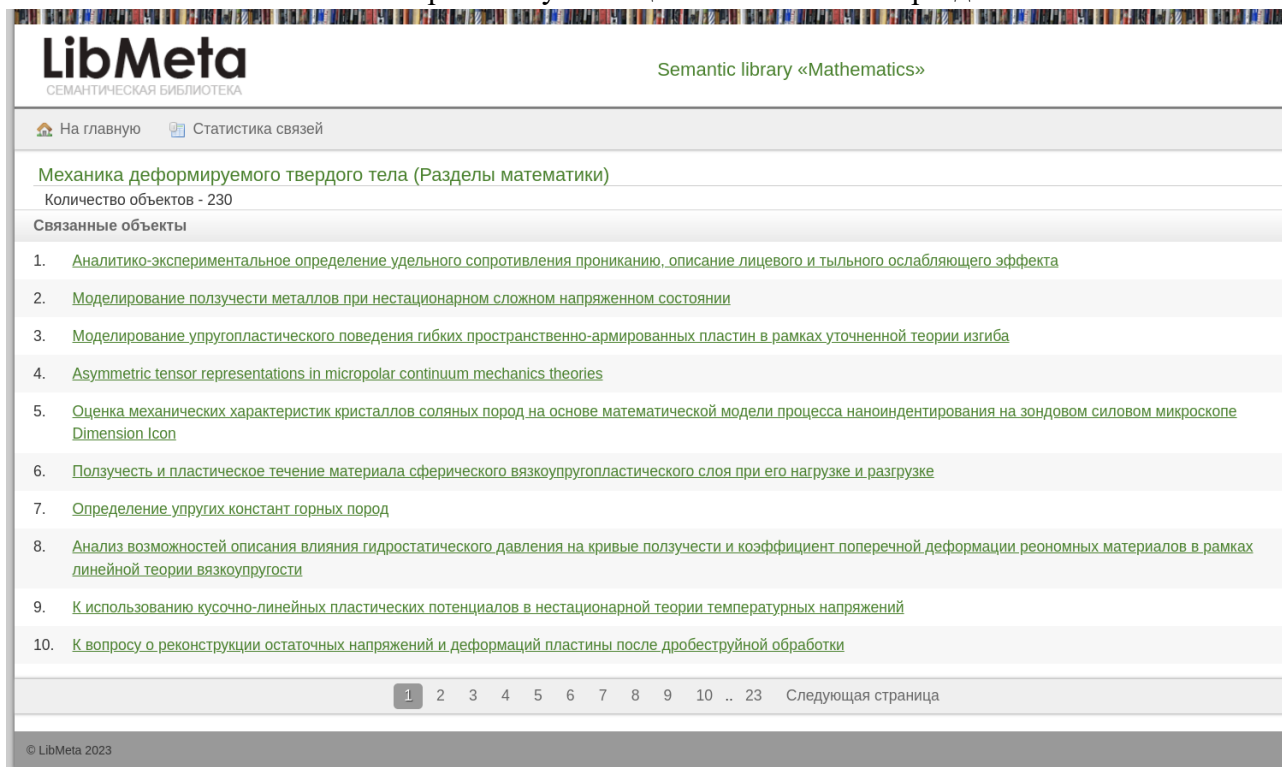


Рис. 11. Скрин 2. Рубрика журнала и ее публикации

В процессе исследований было получено следующее:

- анализ паспортов специальностей позволил выявить наличие общей лексики для различных рубрик журнала, а именно, описание *методов, решений, задач, результатов*;
- анализ контента журнала позволяет *наполнить эти термины «смыслом»*, характерным для каждой специальности конкретно в журнале;
- для тематик журнала были выявлены общие *типы задач*, которые в дальнейшем помогут более глубоко анализировать текст.

Последовательно были реализованы следующие этапы исследований.

Установлены и использованы четыре меры смыслового соответствия:

- эквивалентность рубрик по содержанию,
- включение рубрики ВАК в сопоставленную рубрику журнала,
- включение сопоставленной рубрики журнала в рубрику ВАК,
- значительное пересечение объёмов рубрик журнала.

Решены следующие задачи:

- анализ текстов статей и паспортов ВАК,
- извлечение сущностей и связей,
- поиск УДК,
- поиск схожих
 - персон,
 - публикаций,
- статистика,
- классификация по тематикам,
 - по специальностям,
 - по рубрикам,
- обзоры и сравнения статей.

На рисунке 12 представлен узел графа знаний «Публикация». Пунктиром выделены «стандартные» метаданные публикации. Сплошными стрелками указываются дополнительные связи, которые появились благодаря анализу рубрик и специальностей ВАК.

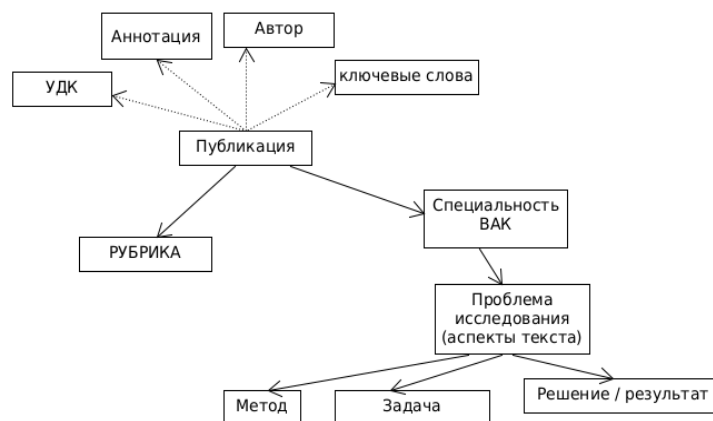


Рис.12. Узел ГЗ «Публикация» с новыми связями

Предложенная в статье методика построения графа знаний включает этапы подготовки данных, выбора методов, анализа результатов, достраивания онтологии. Целью этих этапов является выделение, структуризация и связывание информации по предметной области из сырых текстов, фиксируя ее в виде графа знаний предметной области.

Заключение. Навигация по различным узлам ГЗ позволяет осуществлять поиск, который отличается от классического, реализуемого в библиографических базах данных.

Применение методов обработки естественного языка к графам знаний журнала применимо для насыщения и развития ПрО журнала.

Получаем в перспективе:

- статистику для рецензирования и других целей;
 - соответствие терминов ПрО, ВАК, рубрикам;
 - подбор УДК и других классификаторов;
- и другие элементы анализа статей журнала

Интеллектуальный анализ – единственный инструмент установления соответствия рубрик, создания онтологии разметки и ГЗ.

ПрО журнала, таким образом, будет представлена *задачами, методами, результатами и т.д.*, которым посвящены публикации, при этом вершинами графа ПрО будут рубрики и специальности ВАК.

Перемещаясь по вершинам графа, получаем множества публикаций, которые относятся к рубрике, и в которых изучается та или иная задача, относящаяся к определенной специальности ВАК. Полученный информационный ресурс отличается от того, как журнал представлен в интернете на основе библиографических данных.

Результат позволяет пользователю оценить, *нужна ли конкретная статья или нет, не скачивая, не читая целиком.*

Благодарности. Авторы выражают благодарность редакторскому коллективу журнала «Информационные и математические технологии в науке и управлении» за внимание к предмету исследования статьи, обсуждение и экспертный анализ результатов.

Работа представлена в рамках выполнения темы НИР «Математические методы анализа данных и прогнозирования» ФИЦ ИУ РАН, а также в рамках проекта ИСЭМ СО РАН, № темы FNEU-2021-0007, рег. № АААА-А21-121012090007-7.

Список источников

1. Tryon C. Cluster analysis. London. Ann Arbor Edwards Bros, 1939, 139 p.
2. Oyewole G.J., Thopil G.A. Data clustering: application and trends. *Artif Intell Rev*, 2023, 56, pp.6439–6475, DOI:10.1007/s10462-022-10325-y
3. Pitafi S., Anwar T., Sharif Z. A taxonomy of machine learning clustering algorithms, challenges, and future realms. *Appl. Sci.*, 2023, 13, 3529, DOI: 10.3390/app13063529.
4. Xu Q., Gu H, Ji S. Text clustering based on pre-trained models and autoencoders. *Front. Comput. Neurosci.*, 2024, 17, 1334436, DOI:10.3389/fncom.2023.1334436.
5. Alkaissi H., McFarlane S.I. Artificial Hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, 2023, 15(2): e35179, DOI:10.7759/cureus.35179.
6. Probiez B., Kozak J., Hrabia A., Clustering of scientific articles using natural language processing. *Procedia computer science*, 2022, v. 207, 2022, pp. 3449-3458, DOI:10.1016/j.procs.2022.09.403.
7. Salton G., Wong A., Yang C.S. A vector space model for automatic indexing. 1975. *Communications of the ACM*, vol. 18, pp. 613-620, DOI:10.1145/361219.361220.
8. Mikolov T., Sutskever I., Chen K., et al. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, ArXiv, 2013, DOI:10.48550/arXiv.1310.4546
9. Mikolov T. Efficient estimation of word representations in vector space. 1st Int. Conf. Learn. Represent. ICLR 2013, Work. Track Proc., 2013, p. 1–12.
10. Turian J., Ratinov L., Bengio Y. Word representations: A simple and general method for semi-supervised learning. In *Proceedings of the 48th Annual meeting of the association for computational linguistics, ACL '10*. Association for computational linguistics, 2010.
11. Воронцов К.В. Аддитивная регуляризация тематических моделей коллекций текстовых документов / К.В. Воронцов // Доклады РАН, 2014. – Т. 456. – № 3. – С. 268–271.
12. ruSciBench – бенчмарк для оценки эмбедингов научных текстов. – URL: <https://habr.com/ru/articles/781032/>.
13. Huggingface. Available at: <https://huggingface.co/mlsa-iai-msu-lab/sci-rus-tiny>.
14. Alkaissi H., McFarlane S.I. Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*, 2023, 15(2): e35179, DOI:10.7759/cureus.35179.

15. Cadeddu A., Chessa A., De Leo V. Vincenzo, et al. A comparative analysis of knowledge injection strategies for large language models in the domain. Engineering applications of Artificial Intelligence, 2024, vol.133, part B. 108166, DOI:10.1016/j.engappai.2024.108166.
16. Peng C., Xia F., Naseriparsa M. Osborne F. Knowledge graphs: opportunities and challenges. Available at: <https://arxiv.org/pdf/2303.13948>.
17. Ataeva O., Serebryakov V., Tuchkova N. Ontological approach to a knowledge graph construction in a semantic library. Lobachevskii journal of mathematics, 2023, vol. 44, no. 6, pp. 2229–2239, DOI:10.1134/S1995080223060471.
18. Ataeva O., Kornet Yu.N., Serebryakov V., Tuchkova N., Approach to creating a thesaurus and a knowledge graph of an applied subject area. Lobachevskii journal of mathematics, 2023, vol. 44, no. 7, pp. 2577–2586.
19. Blashfield R.K. The growth of cluster analysis: Tryon, Ward, And Johnson, Multivariate behavioral research, 1980, 15:4, pp. 439-458, DOI:10.1207/s15327906mbr1504_4.
20. Blei D.M. Probabilistic topic models. Communications of the ACM, 2012, vol. 55, no. 4, pp. 77–84.
21. Blei D.M., A.Y.Ng, Jordan M.I. Latent dirichlet allocation. Journal of machine learning research, 2003, 3, pp. 993-1022.
22. Ikotun A.M., Ezugwu A.E., Abualigah L., et al. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information sciences, 2023, vol. 622, pp. 178-210, DOI:10.1016/j.ins.2022.11.139, available at: <https://www.sciencedirect.com/science/article/pii/S0020025522014633> (accessed: 10/13/2024)
23. Sparck J.K. A statistical interpretation of term specificity and its application in retrieval. Journal of documentation, 1972, vol. 28, no. 1, pp. 11-21, DOI:10.1108/eb026526.
24. Chowdhary K.R. Natural language processing. Fundamentals of artificial intelligence, 2020, pp. 603-649.
25. Young T. et al. Recent trends in deep learning based natural language processing. IEEE Computational intelligence magazine, 2018, vol. 13, no. 3, pp. 55-75.

Атаева Ольга Муратовна. Старший научный сотрудник Вычислительного центра им. А.А. Дородницына ФИЦ ИУ РАН, кандидат техн. наук, специалист в области системного программирования и баз данных. ORCID: 0000-0003-0367-5575, oli.ataeva@gmail.com, 119333, Россия, Москва, ул. Вавилова, д.40

Массель Людмила Васильевна. Специалист в области информационных технологий и искусственного интеллекта, главный научный сотрудник ИСЭМ СО РАН, зав. отделом «Системы искусственного интеллекта в энергетике, главный редактор журнала «Информационные и математические технологии в науке и управлении», доктор техн. наук, профессор, заслуженный деятель науки РФ, ORCID: 0000-0002-9088-9012, masse1@isem.irk.ru, 664033, Россия, Иркутск, Лермонтова, 130

Серебряков Владимир Алексеевич. Специалист в области теории формальных языков и её приложений, доктор физ.-мат. наук, профессор, зав. отделом Вычислительного центра им. А.А. Дородницына ФИЦ ИУ РАН. Руководитель и участник разработки ряда известных программных проектов, в частности ИСИР и ИСИР РАН, Научный портал РАН, ORCID: 0000-0003-1423-621X, serebrvas@gmail.com, 119333, Россия, Москва, ул. Вавилова, д.40

Тучкова Наталия Павловна. Старший научный сотрудник Вычислительного центра им. А.А. Дородницына ФИЦ ИУ РАН, кандидат физ.-мат. наук, специалист в области алгоритмических языков и информационных технологий, ORCID: 0000-0001-5357-9640, tuchkova.nataly@gmail.com, 119333, Россия, Москва, ул. Вавилова, д.40

UDC 004.8+519.217 + 519.876.2

DOI:10.25729/ESI.2024.35.3.001

Data mining when constructing a knowledge graph of a multidisciplinary journal

Olga M. Ataeva¹, Ludmila V. Massel², Vladimir A. Serebryakov¹, Natalia P. Tuchkova¹

¹FRC CSC RAS, Russia, Moscow, oli.ataeva@gmail.com

²Melentiev Energy Systems Institute SB RAS, Russia, Irkutsk

Abstract. The paper explores the thematic diversity of the interdisciplinary journal. The purpose of the research is to build a knowledge graph of the journal for the thematic presentation and systematization of the electronic

archive and new publications of the journal. The initial data are journal articles devoted to various information and mathematical technologies in science and management, that is, interdisciplinary research. The systematization of texts using vector analysis methods is proposed. In the process of thematic analysis of the content of the journal, a division into headings is proposed, links of headings and articles with the corresponding descriptions of the specialties of the Higher Attestation Commission are established. To analyze the topic, an exploratory analysis of the source texts is used, then data mining methods are used. The results of the division are provided to the experts of the journal, after which a decision is made on the formation of a thematic heading and the inclusion of the specialties of the Higher Attestation Commission in it. The journal articles are integrated into the LibMeta semantic library, which is why the library's ontology is being completed and the journal's ontology is being formed, and the journal's knowledge graph is being built on this basis. A procedure for navigating through the content of the journal using the knowledge graph in the LibMeta semantic library is proposed, which can become the basis for information support of scientific research and the creation of a digital assistant in an interdisciplinary subject area. Examples are given for specific journal content, but the proposed technology can be extended to other journals, since most journals belonging to several specialties of the Higher Attestation Commission naturally capture several disciplines.

Keywords: knowledge graph, semantic library, ontology completion, clustering of scientific articles, text summarization

Acknowledgements: The work is presented within the framework of the research topic "Mathematical Methods of Data Analysis and Forecasting" of the FRC ICS RAS, as well as within the framework of the ISEM SB RAS project, topic No. FNEU-2021-0007, reg. No. AAAA-A21-121012090007-7.

References

1. Tryon C. Cluster analysis. London. Ann Arbor Edwards Bros, 1939, 139 p.
2. Oyewole G.J., Thopil G.A. Data clustering: application and trends. *Artif Intell Rev*, 2023, 56, pp.6439–6475, DOI:10.1007/s10462-022-10325-y
3. Pitafi S., Anwar T., Sharif Z. A taxonomy of machine learning clustering algorithms, challenges, and future realms. *Appl. Sci.*, 2023, 13, 3529, DOI: 10.3390/app13063529.
4. Xu Q., Gu H, Ji S. Text clustering based on pre-trained models and autoencoders. *Front. Comput. Neurosci*, 2024, 17, 1334436, DOI:10.3389/fncom.2023.1334436.
5. Alkaissi H., McFarlane S.I. Artificial Hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, 2023, 15(2): e35179, DOI:10.7759/cureus.35179.
6. Probiez B., Kozak J., Hrabia A., Clustering of scientific articles using natural language processing. *Procedia computer science*, 2022, v. 207, 2022, pp. 3449-3458, DOI:10.1016/j.procs.2022.09.403.
7. Salton G., Wong A., Yang C.S. A vector space model for automatic indexing. 1975. *Communications of the ACM*, vol. 18, pp. 613-620, DOI:10.1145/361219.361220.
8. Mikolov T., Sutskever I., Chen K., et al. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, ArXiv, 2013, DOI:10.48550/arXiv.1310.4546
9. Mikolov T. Efficient estimation of word representations in vector space. 1st Int. Conf. Learn. Represent. ICLR 2013, Work. Track Proc., 2013, p. 1–12.
10. Turian J., Ratinov L., Bengio Y. Word representations: A simple and general method for semi-supervised learning. In *Proceedings of the 48th Annual meeting of the association for computational linguistics, ACL '10*. Association for computational linguistics, 2010.
11. Vorontsov K.V. Additivnaya regularizatsiya tematicheskikh modeley kollektiy tekstovoykh dokumentov [Additive regularization of topic models of text document collections]. *Doklady RAN [Reports of the RAS]*, 2014, vol. 456, no.3, pp. 268–271.
12. ruSciBench – бенчмарк для оценки эмбедингов научных текстов. – URL: <https://habr.com/ru/articles/781032/>.
13. Huggingface. Available at: <https://huggingface.co/mlsa-iai-msu-lab/sci-rus-tiny>.
14. Alkaissi H., McFarlane S.I. Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*, 2023, 15(2): e35179, DOI:10.7759/cureus.35179.
15. Cadeddu A., Chessa A., De Leo V. Vincenzo, et al. A comparative analysis of knowledge injection strategies for large language models in the domain. *Engineering applications of Artificial Intelligence*, 2024, vol.133, part B. 108166, DOI:10.1016/j.engappai.2024.108166.
16. Peng C., Xia F., Naseriparsa M. Osborne F. Knowledge graphs: opportunities and challenges. Available at: <https://arxiv.org/pdf/2303.13948>.
17. Ataeva O., Serebryakov V., Tuchkova N. Ontological approach to a knowledge graph construction in a semantic library. *Lobachevskii journal of mathematics*, 2023, vol. 44, no. 6, pp. 2229–2239, DOI:10.1134/S1995080223060471.

18. Ataeva O., Kornet Yu.N., Serebryakov V., Tuchkova N., Approach to creating a thesaurus and a knowledge graph of an applied subject area. Lobachevskii journal of mathematics, 2023, vol. 44 , no. 7, pp. 2577–2586.
19. Blashfield R.K. The growth of cluster analysis: Tryon, Ward, And Johnson, Multivariate behavioral research, 1980, 15:4, pp. 439-458, DOI:10.1207/s15327906mbr1504_4.
20. Blei D.M. Probabilistic topic models. Communications of the ACM, 2012, vol. 55, no. 4, pp. 77–84.
21. Blei D.M., A.Y.Ng, Jordan M.I. Latent dirichlet allocation. Journal of machine learning research, 2003, 3, pp. 993-1022.
22. Ikotun A.M., Ezugwu A.E., Abualigah L., et al. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information sciences, 2023, vol. 622, pp. 178-210, DOI:10.1016/j.ins.2022.11.139, available at: <https://www.sciencedirect.com/science/article/pii/S0020025522014633> (accessed: 10/13/2024)
23. Sparck J.K. A statistical interpretation of term specificity and its application in retrieval. Journal of documentation, 1972, vol. 28, no. 1, pp. 11-21, DOI:10.1108/eb026526.
24. Chowdhary K.R. Natural language processing. Fundamentals of artificial intelligence, 2020, pp. 603-649.
25. Young T. et al. Recent trends in deep learning based natural language processing. IEEE Computational intelligence magazine, 2018, vol. 13, no. 3, pp. 55-75.

Ataeva Olga Muratovna. Senior researcher of the of Dorodnicyn computing center FRC SCS RAS, PhD, expert in the field of system programming and databases, oli.ataeva@gmail.com, ORCID: 0000-0003-0367-5575, 119333, Russia, Moscow, Vavilova, 40.

Massel Liudmila Vasilievna. Expert in the field of information technologies and artificial intelligence, chief researcher of the ESI SB RAS, head of the department of "Artificial Intelligence Systems in Energy sector", editor-in-chief of the journal "Information and mathematical technologies in science and management", Doctor of technical sciences, professor, honored scientist of the Russian Federation, massel@isem.irk.ru, ORCID: 0000-0002-9088-9012, 664033, Russia, Irkutsk, Lermontov, 130.

Serebryakov Vladimir Alekseevich. Expert in the field of theory of formal languages and its applications, doctor of sciences, professor, head of Dorodnicyn computing center FRC SCS RAS department, head and participant in the development of a number of well-known program projects, in particular ISIR and ISIR RAS, scientific portal RAS, ORCID: 0000-0003-1423-621X, serebrvas@gmail.com, 119333, Russia, Moscow, Vavilova, 40.

Tuchkova Natalia Pavlovna. Senior researcher of Dorodnicyn computing center FRC SCS RAS, PhD in physics with a math degree, graduated from CS Faculty of Lomonosov MSU, the expert in the field of algorithmic languages and information technologies, ORCID: 0000-0001-5357-9640, tuchkova.nataly@gmail.com, 119333, Russia, Moscow, Vavilova, 40.

Статья поступила в редакцию 30.07.2024; одобрена после рецензирования 11.10.2024; принята к публикации 14.10.2024.

The article was submitted 07/30/2024; approved after reviewing 10/11/2024; accepted for publication 10/14/2024.

Онтологический анализ приемов аргументации в научном дискурсе

Сидорова Елена Анатольевна, Кононенко Ирина Семеновна

Институт систем информатики им. А.П. Ершова СО РАН,
Россия, Новосибирск, *lsidorova@iis.nsk.su*

Аннотация. Статья посвящена анализу дискурсивных приёмов аргументации в рамках современных тенденций в области моделирования рассуждений. Рассуждение в научном дискурсе организуется, как последовательность дискурсивных приемов, т.е. шагов рассуждения, соответствующих определенным мыслительным операциям над объектом предметной области. С формальной точки зрения элементарной структуре рассуждения соответствует аргумент, который представляет собой переход от посылок к доказываемому тезису. В работе представлена онтологическая модель, описывающая типовые аргументы – схемы аргументации, основанная на теории Дугласа Уолтона. Обсуждаются вопросы классификации существующих способов рассуждения и рассматривается подход к онтологическому анализу структур дискурса на основе онтолого-семантических отношений, которые обеспечивают концептуальную основу для моделей рассуждения и помогают понять природу сущностей, отношений и рассуждений, связанных с конкретным контекстом. В качестве материала для практического исследования использовался аннотированный корпус текстов, относящихся к области научной коммуникации, включающий аналитические статьи с комментариями пользователей с форума «Хабр» (научно-популярный дискурс) и научные статьи с комментариями рецензентов (собственно научный дискурс). Особый интерес в данном исследовании представляли конфликтные аргументативные приемы, которые используются в текстах исследуемых жанров: Антитезис (‘атака’+‘на_тезис’) и Контраргумент (‘атака’+‘на_аргумент’). Проведены статистический анализ и дальнейшая классификация этих приемов в зависимости от жанра. Полученная классификация основана на 12 типовых приемах первого типа и 30 приемах второго типа. Онтология, классификация приёмов и корпус размещены в открытом доступе на платформе ArgNetBank Studio.

Ключевые слова: анализ аргументации, онтология аргументации, схема аргументации, полемические приемы

Цитирование: Сидорова Е.А. Онтологический анализ приемов аргументации в научном дискурсе / Е.А. Сидорова, И.С. Кононенко // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 20-32. – DOI: 10.25729/ESI.2024.35.3.002.

Введение. Научный дискурс характеризуется наличием ряда выдвигаемых автором утверждений, которые рассматриваются и обосновываются в тексте с различных точек зрения. При этом автор не только доказывает свою позицию путем логических рассуждений, но и мысленно дискутирует с возможным оппонентом, моделируя его контрдоводы. Дискурсивная стратегия организуется, как последовательность дискурсивных приёмов, или шагов рассуждения, соответствующих определенным мыслительным операциям над объектом предметной области. С формальной точки зрения, элементарной структуре рассуждения соответствует аргумент, который представляет собой переход от посылок к доказываемому или опровергаемому тезису, а всё множество рассуждений соответствует понятию аргументации, т.е. приведению доводов (аргументов) с целью изменения или формирования некоторого убеждения (позиции) у читателя или слушателя.

Существуют различные подходы к моделированию аргументации [1]. На верхнем уровне выделяют два типа аргументации: абстрактную и структурированную [2]. Абстрактная аргументация рассматривает каждый аргумент, как атомную сущность без внутренней структуры, в то время, как структурированная аргументация принимает во внимание внутреннюю структуру каждого аргумента. Одной из самых влиятельных моделей аргументации, относящихся к типу структурированной аргументации, является модель Стивена Тулмина [3]. Согласно этой модели, аргумент в общем случае включает шесть компонентов: тезис, посылки, обоснование

вывода, поддержка утверждений, границы применимости и степень уверенности. Многие модели аргументации в той или иной степени опираются на приведённую модель Тулмина.

Развитие моделей аргументации тесно связано с практическим применением в компьютерных науках. Наиболее известной моделью, нашедшей применение в практических системах анализа аргументации, является модель Дугласа Уолтона [4], в которой структурированный аргумент определяется, как набор утверждений, состоящий из трёх частей: посылок, заключения (тезиса) и вывода из посылок к заключению. Уолтон ввёл понятие схемы аргумента – формы, представляющей стереотипную модель рассуждения. Такие схемы также называют моделями правдоподобного вывода, в том смысле, что если посылки аргумента верны, то предположительно верно и заключение [5]. В работе [6] приведён компендиум, содержащий 60 основных схем аргументации.

Однако практическое применение данного набора осложнено следующими факторами:

- многие схемы имеют варианты (подсхемы);
- некоторые схемы не включены в компендиум и обсуждаются отдельно;
- некоторые схемы похожи, и в конкретной ситуации выбор не всегда очевиден;
- иногда в качестве утверждения (посылки или заключения) может выступать другой аргумент;
- множество схем является открытым, это означает, что при рассмотрении новых текстовых жанров может возникать потребность введения новых схем.

Указанные осложнения приводят к неоднозначности при аннотировании текстов, разногласию между разметчиками, ухудшению качества создаваемых наборов данных для применения методов машинного обучения и, как следствие, низкому качеству автоматизации анализа аргументации в целом. Таким образом, назрела необходимость более глубокого анализа и систематизации уже существующих моделей аргументов.

Можно выделить две основные задачи, необходимые для решения этих проблем. Во-первых, формальное представление множества схем с использованием формального языка, например, OWL – языка построения онтологий, – для строгого описания взаимосвязей компонентов аргументации с дальнейшим подключением поисковых сервисов [7]. Другая задача связана с построением классификаторов, позволяющих объединять модели или схемы аргументов в группы по значимым признакам. Такие классификаторы помогали бы как разметчикам при определении типа аргумента и выработке методологии разметки, так и разработчикам интеллектуальных систем анализа аргументации.

Работ, посвященных данным вопросам, немного. Это, по-видимому, связано со сложностью внедрения новых моделей в практику. Большинство существующих наборов данных (датасетов) используют базовую модель аргументации без выделения типов аргументов, т.е. размечается только связь между утверждениями – поддержка или атака (ArgMicro [8, 9], русскоязычный корпус для соревнования RuARG-2022 [10], корпуса для анализа дебатов и обсуждений [11]). Иногда выделяется ограниченное число типов аргументации, как в работах [12, 13]. Для разметки предлагается таксономическая иерархия схем аргументации, основанная на аргументационной модели топосов (Argumentum Model of Topics) с ограниченным количеством схем (8 схем среднего уровня), выделенных на смысловых, взаимоисключающих отличительных критериях и применимых к широкому спектру контекстов [14]. С другой стороны, разработчики систем автоматического анализа аргументации также ограничиваются небольшим количеством наиболее распространенных типов распознаваемых аргументов или упрощенной моделью, например, в [15-18]. Обзор работ, посвященных подготовке датасетов, методам и подходам в области автоматического анализа аргументации, приводится в [19].

Одним из самых известных исследований, связанным с систематизацией аргументов, является Периодическая таблица аргументов, предложенная Жаном Вагемансом [20]. В данной таблице представлена категоризация аргументов, которая учитывает как форму, так и содержание аргументов, что позволяет, по задумке авторов, упростить выбор аргумента при аннотировании текста. На сайте проекта (<https://periodic-table-of-arguments.org>) приведены 17 типов аргументов, расположенных в четырех квадрантах таблицы, в зависимости от типа субъекта и предиката в посылках и заключении [21]. В работе [22] авторы описывают опыт параллельного аннотирования корпуса текстов с использованием двух типологий, демонстрирующих различные теоретические перспективы и обоснования: таксономия схем аргументации Дугласа Уолтона представляет собой “эмпирически ориентированную классификацию схем, основанную на изучении очевидных конвенций аргументативной практики” и периодическая таблица аргументов Жана Вагеманса, которая “находится на другом конце спектра, начиная с множественных априорных критериев, разработанных для исчерпывающего описания всех возможных реализаций в реальной практике”.

Целью предлагаемой статьи является анализ приёмов аргументации, представленных таксономией Дугласа Уолтона, и их систематизация в рамках онтологии аргументации. Исследования проводились на корпусе русскоязычных текстов научной коммуникации. Дальнейшая структура статьи отражает основные этапы и методику проведенного исследования. В разделе 1 приведена формализация области аргументации, полученная методом онтологического моделирования. В разделе 2 предложена общая классификация типов аргументов и проведен частотный анализ выделенных категориальных признаков. Раздел 3 посвящен сравнительному анализу приёмов аргументации на основе корпусного исследования текстов, относящихся к разным жанрам научной коммуникации. В разделе 4 обсуждаются результаты исследования.

1. Моделирование аргументации. Визуально множество аргументов обычно представляется в виде ориентированного графа в соответствии с форматом AIF (Argument Interchange Format), в котором выделяются два типа вершин: информационные вершины – утверждения *Inode* и аргументы *SNode*. В данной работе для исследования разных типов аргументов используется онтология аргументации [23], основанная на AIF-онтологии [24], включающей описание схем аргументов из компендиума Д. Уолтона.

В разрабатываемой онтологии можно выделить следующие компоненты:

$$O = \langle O^G, O^S, O^C, O^F, ST, A, Sim \rangle, \quad (1)$$

где $O^G = \langle SNode, Inode, R^N \rangle$ – графовое представление аргументации, содержащее вершины двух типов *SNode*, *Inode* и бинарные отношения между вершинами R^N ;

$O^S = O^{SA} \cup O^{ST}$ – онтология предметной области аргументации, содержащая классы типовых схем рассуждения O^{SA} и классы утверждений O^{ST} ; классы типовых схем (далее схемы) формально описывают внутреннюю структуру аргументов – экземпляров схем;

$O^C = \{O^{CT}, O^{CP}, O^{CS}, O^{CA}\}$ – классификатор типовых схем аргументации, включающий четыре компонента, позволяющих группировать схемы по четырём различным основаниям: O^{CT} – на основе типа тезиса, O^{CP} – на основе типа посылки, O^{CS} – на основе семантического отношения и O^{CA} – на основе типа атаки;

O^F – онтология метаописаний, содержащая классы дескрипторов (форм) для схем аргументации и их атрибутов. Эта онтология позволяет задавать языковые описания схем аргументов (см. рис.1);

$ST = \{st \in T | T \in O^{ST}\}$ – множество аргументативных утверждений;

$A = \{a \in S | S \in O^{SA}\}$ – множество аргументов;

Sim – отношение эквивалентности между утверждениями; данное отношение связывает семантически близкие утверждения, либо утверждение-отсылку с его референтом.

Исходная АИФ-онтология была модифицирована нами следующим образом: введена многоаспектная классификация схем аргументации; добавлен ряд схем, необходимых для анализа научной коммуникации; введено отношение эквивалентности, что позволяет обеспечить лучшую связность графа аргументации при аннотировании текста.

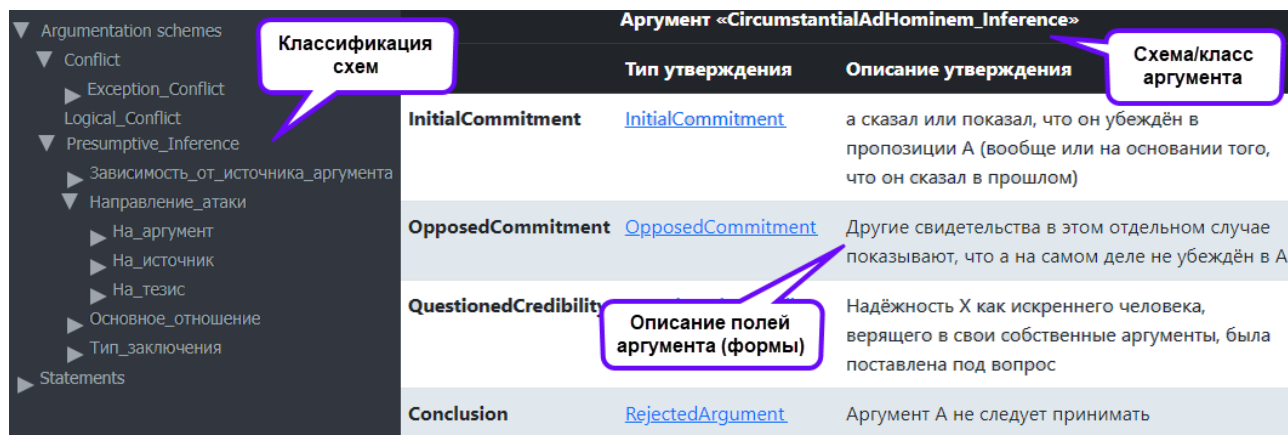


Рис. 1. Онтологическое представление полемической схемы аргументации

Для анализа и аннотирования текстов создана платформа ArgNetBank Studio (<https://uniserv.iis.nsk.su/arg>) [25], на которой представлены онтология аргументации и описание схем, используемых при аннотировании текстов (см. рисунок 1). С помощью данной платформы был создан аннотированный корпус текстов, включающий 286 текстов, относящихся к сфере научной коммуникации, в том числе тексты с комментариями. Тексты имеют в среднем объём 3,5 тыс. токенов, средний объём комментария к ним – 2-5 предложений. Всего разметка содержит около 13 тыс. размеченных аргументов.

2. Классификация моделей рассуждений. В [26] была предложена многоаспектная классификация схем аргументации. В данной работе мы развиваем и более подробно рассматриваем полемическую аргументацию. В основе дискурсивного приёма лежит онтолого-семантическое отношение как концептуальная база рассуждения и доказательства в конкретном контексте. Рассматриваются 8 типов онтолого-семантических отношений, наиболее характерных для рассуждений в научных текстах: Каузальность, Гипер-гипонимия, Меронимия, Аналогия, Корреляция, Способ, Коммуникация и Противоречие.

В таблице 1 представлены используемые на текущий момент отношения и их статистическая оценка на корпусе текстов: F_s – количество схем, соотнесённых с каждым из отношений, F_a – количество аргументов данного типа.

Таблица 1. Классификация на основе онтолого-семантического отношения

Отношение	F_s	F_a	Отношение	F_s	F_a
Каузальность	19	5,5 тыс.	Способ	5	1,5 тыс.
Гипер-гипонимия	5	2,6 тыс.	Корреляция	2	1,3 тыс.
Меронимия	4	1,7 тыс.	Коммуникация	5	1,3 тыс.
Противоречие	32	1,7 тыс.	Аналогия	4	200

Отношению Противоречия соответствуют полемические структуры, которым в работе [27] сопоставлены три класса конфликтных отношений:

- опровержение – атака на заключение аргумента,
- подрыв – атака на посылку существующего аргумента,
- подрез – атака на отношение вывода.

Возможность типизации полемических аргументативных ходов предопределяется исходной онтологией, в которой, с одной стороны, выделяется специальный класс конфликтных схем, которые реализуют посылки-исключения, заложенные в описании соответствующих им

аргументов, с другой стороны, во множестве стандартных схем можно выделить подмножество атакующих, которые классифицируются по признаку направления атаки. Среди конфликтных схем можно выделить отражающие схемы (‘отражение атаки’), которые реализуют посылки-исключения стандартных атакующих схем (Таблица 2).

Таблица 2. Классификация полемических аргументов

Атакующие схемы	F_s	F_a	Конфликты	F_s	F_a
На тезис	11	594	На тезис	1	826
На аргумент	4	150	На аргумент	16	368
На источник	4	23	Отражение атаки	6	2

Множество атакующих схем подразделяются на Антитезис (‘атака’ + ‘тезис’) и Контраргумент, который либо непосредственно атакует отношение вывода (‘атака’ + ‘аргумент’), либо опосредованно через посылку (‘атака’ + ‘источник’).

Антитезисы отрицают справедливость тезиса:

- От Негативных Последствий (действие не следует предпринимать ввиду негативных последствий),
- Фальсификация Гипотезы (гипотеза неверна ввиду ложности доказательства).
- Логический Конфликт (утверждение ложно).

Контраргументы атакуют аргумент в целом:

- От Непоследовательности Убедений (аргумент не может быть принят ввиду противоречивости высказанных утверждений).
- От Неясности Вербальной Классификации (аргумент, появившийся в контексте диалога, следует отвергнуть как несостоятельный ввиду того, что некоторое свойство, встречающееся в аргументе, определено слишком неясно, а диалог требует определенной степени точности).

К конфликтному типу рассуждений относятся следующие:

- Альтернативные цели (имеются альтернативные цели помимо выбранных).
- Предвзятость классификации (аргумент основан на условном и предвзятом определении, которое сомнительно).
- Альтернативные средства (имеются альтернативные средства достижения цели, помимо выбранных).
- Сомнение в общепринятом (есть серьезные основания для сомнения в истинности общепринятого мнения).
- Недостаток надёжности (источник информации не является честным).
- Знак других событий (имеются другие события, более надёжно объясняющие знак).

К отражающему типу рассуждений относятся следующие:

- Апелляция к иррациональному страху (в данном случае атака апеллирует к иррациональному страху).
- Ложная дилемма (указанная альтернатива не является дилеммой).

Всего в текущей версии онтологии представлено 42 полемических схемы.

3. Полемические приёмы в научной коммуникации. Убеждение целевой аудитории в справедливости высказываемых в тексте тезисов основано на применении тех или иных приёмов аргументации. Монологический текст опирается на представления автора об очень общей целевой аудитории, в то время как для диалогического характерны более конкретные представления участников общения друг о друге. В отличие от монолога, диалог нацелен не только на передачу информации целевой аудитории, но и на получение от неё обратной связи.

3.1. Polemicheskie zhany. В данной работе особое внимание уделяется жанрам научных статей с комментариями рецензентов и научно-аналитических статей с комментариями-дискуссиями по поводу их содержания (Хабр-форум: <https://habr.com/>). Выбранные жанры обладают свойством нацеленности на получение обратной связи. В таблице 3 проведено сравнение выбранных жанров по значимым критериям.

Таблица 3. Сравнение полемических жанров

Признаки	Статьи с комментариями рецензента	Статьи с комментариями-дискуссиями
<i>Количество участников</i>	2 участника	Не ограничено
<i>Тип обмена репликами</i>	Компонентный диалог	Каскадный диалог
<i>Статус участников</i>	Реагирующий агент-рецензент более авторитетен	Равенство всех участников по статусу
<i>Иерархия коммуникативных задач</i>	Уточнение аспектов, совершенствование текста, одностороннее общение	Обмен мнениями, общение

Научная статья с комментариями рецензента (далее *Ср*) представляет собой обмен стимулирующей и реагирующей репликами. Компонент этого диалога – однократный обмен репликами. Особенность диалога состоит в том, что инициатором каждого компонента всегда является один и тот же участник – автор статьи, на которую реагирует рецензент. Предопределённый статус эксперта характерен для рецензента в собственно научной сфере коммуникации с имеющейся в ней иерархией.

Общение на Хабр-форуме с помощью комментариев (далее *Ск*) – многосторонняя форма общения, которое не ограничивается двумя участниками и представляет собой обмен репликами, каждая из которых реагирует на саму статью или на некоторую высказанную ранее в связи с этой статьёй реплику, с возможностью смены участников (полилог). Равенство всех участников по статусу характерно для научно-популярной сферы коммуникации.

С точки зрения иерархии задач [28], отличие *Ср* от *Ск* состоит в том, что второй нацелен на получение обратной связи в виде экспликации знаний и мнений, т.е. *Ск* обеспечивает многостороннее общение, а задача *Ср* предполагает одностороннее общение, которое обеспечивает некую прагматическую задачу – создание условий для публикации статьи.

При исследовании жанров с обратной связью особый интерес представляют типовые ходы в комментариях и статистика по ним. Можно предположить, что не последнюю роль в текстах этих жанров играют полемические ходы, которые позволяют рецензенту/комментатору противопоставить точке зрения автора статьи/собеседника собственную позицию по вопросам, которые относятся к тематике или форме изложения статьи.

3.2. Анализ полемических приёмов. Для исследования особенностей полемических жанров потребовалось провести анализ и классификацию возможных в рамках онтологии полемических аргументативных ходов на основе двух жанров размеченного корпуса. Общее количество полемических аргументов в корпусе составило 1359 (*Ср* = 324, *Ск* = 1035).

В таблице 4 представлен статистический анализ распределения элементарных аргументативных ходов, представленных схемами, по жанрам (*Fr* – относительная частота аргументов заданной схемы в корпусе на 10 тыс. аргументов). В таблице представлены 11 самых частотных схем из 42 представленных в корпусе.

Полемические ходы, т.е. простые приёмы и основанные на них составные приёмы, используются как в монологических, так и в диалогических текстах. В монологических (статьях, новостях, рецензиях) их использование базируется на представлениях автора об альтернативных мнениях или подходах к обсуждаемой теме. Несомненно, альтернативные мнения более репрезентативны в комментариях исследуемых жанров диалогического характера (*Ср* и *Ск*).

Поэтому далее будут приведены наиболее типичные примеры на основе комментариев из *Ср* и *Ск*.

Таблица 4. Относительная частота полемических схем в полемических жанрах

Полемическая схема	Тип приёма	<i>Fr (Ср)</i>	<i>Fr (Ск)</i>
Логический конфликт	конфликт, на тезис	745	1023
От неясности вербальной классификации	атака, на аргумент	356	51
Негативные следствия	атака, на тезис	149	327
Противоположные последствия	конфликт, на аргумент	149	72
От произвольной вербальной классификации	атака, на аргумент	109	45
Конфликтные цели	конфликт, на аргумент	103	12
Неактуальность утверждения	конфликт, на аргумент	80	6
Сомнение в общепринятом	конфликт, на аргумент	23	76
Альтернативные средства	конфликт, на аргумент	29	72
Знак других событий	конфликт, на аргумент	6	56
Отличия подрывают сходство	конфликт, на аргумент	0	49

Атакующий приём «Логический Конфликт» приблизительно поровну представлен в обоих жанрах (см. рисунок 2¹). На рисунках 2а, 2б числа под утверждениями указывают на порядковый номер комментария к статье в источнике.



Рис. 2. Приём “Логический конфликт” в разных жанрах

Контраргумент «От Неясности Вербальной Классификации» используется в *Ср* по сравнению с *Ск* на порядок чаще (356 vs. 51). Приведем метаописание данной схемы аргумента.

Посылка 1: Аргумент *A* встречается в контексте диалога, требующего определённого уровня точности.

Посылка 2: Некоторое свойство *F*, встречающееся в аргументе *A*, определено слишком неясно, чтобы удовлетворять требованиям уровня точности, допустимым в контексте диалога.

Заключение: Аргумент *A* не следует принимать.

Из самого толкования этой атакующей схемы видно, что это специфически диалогический приём, когда рецензент с позиции читателя требует от автора полной ясности изложения (см. рисунок 3). В основе подобных полемических ходов в жанре *Ср* лежит прагматическая задача «совершенствования текста» и статус «доминирования» рецензента над автором.

¹ На рисунках 2-4 пометка а) соответствует жанру *Ср*, а пометка б) – жанру *Ск*.

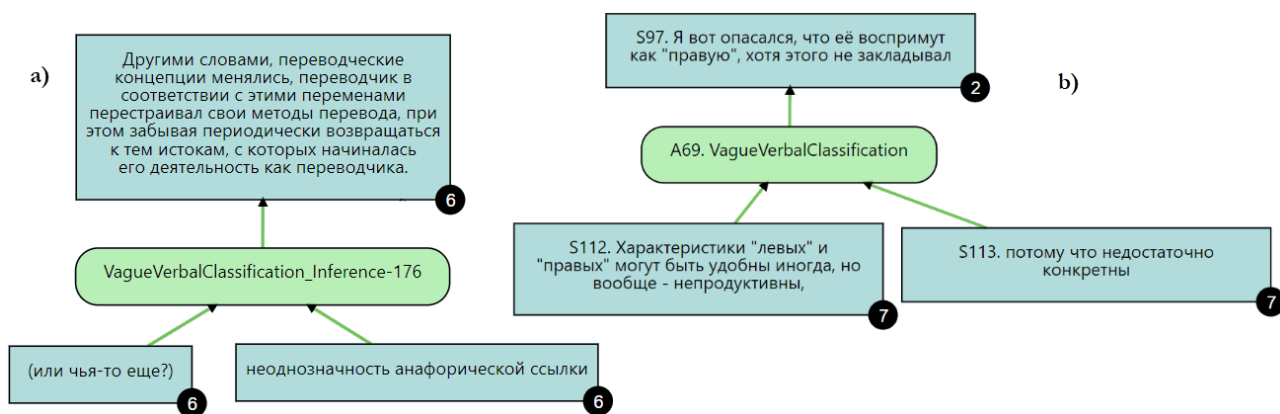


Рис. 3. Приём «От Неясности Вербальной Классификации» в разных жанрах

В обоих жанрах встречается контраргумент «От Произвольной Классификации», близкий к предыдущему в том плане, что это также специфически диалогический приём, с помощью которого оппонент указывает контрагенту на неверно выбранный термин для определения некоторого ключевого свойства в контексте диалога, что означает недостаточность приведенного контрагентом аргумента.

Посылка 1: Аргумент *A* встречается в контексте диалога, который требует точного определения ключевого свойства *F*.

Посылка 2: Некоторое свойство *F*, встречающееся в аргументе *A*, определено произвольным образом.

Заключение: Аргумент *A* не следует принимать.

Примеры данного приёма приведены на рисунке 4.

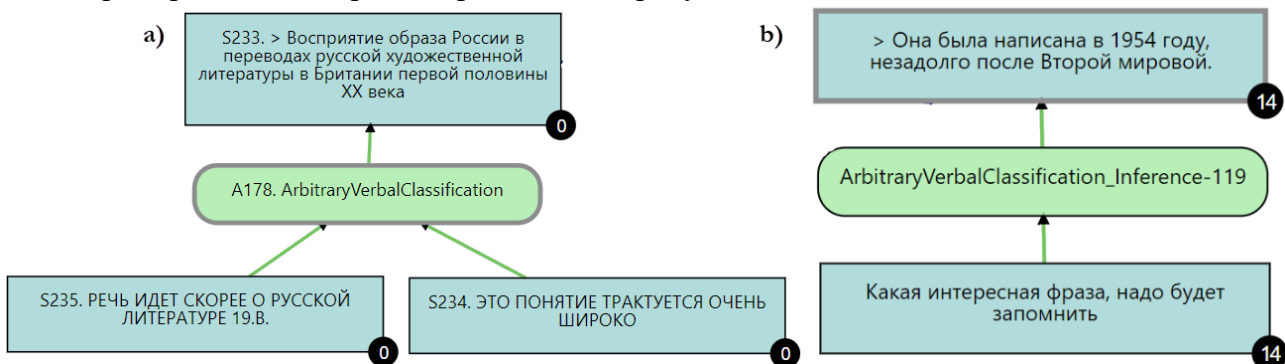


Рис. 4. Приём «От Произвольной Классификации» в разных жанрах

Использование данного приёма в *Ср* в два с лишним раза чаще (109 vs. 45), чем в *Ск*, что не удивительно, однако интересно то, что форум, который не нацелен на обучение/совершенствование реплик партнеров, приводит такую аргументацию с долей иронии.

Существуют простые приёмы, которые имеют нулевую частотность в *Ср* и достаточно представлены в комментариях *Ск*. Речь идет о психологических приёмах, таких, как довод к человеку (обращение к личным качествам или поступкам того лица, чья идея обсуждается). В предложенной классификации это – контраргументы, основанные на свойствах личности:

- «От предвзятости» (19 vs. 0): Предвзятый оппонент не склонен принимать во внимание все факты с обеих сторон (рисунок 5а).
- «Апелляция к личности» (10 vs. 0): Оппонент – человек с плохим характером, поэтому его аргументы не принимаются (рисунок 5б).

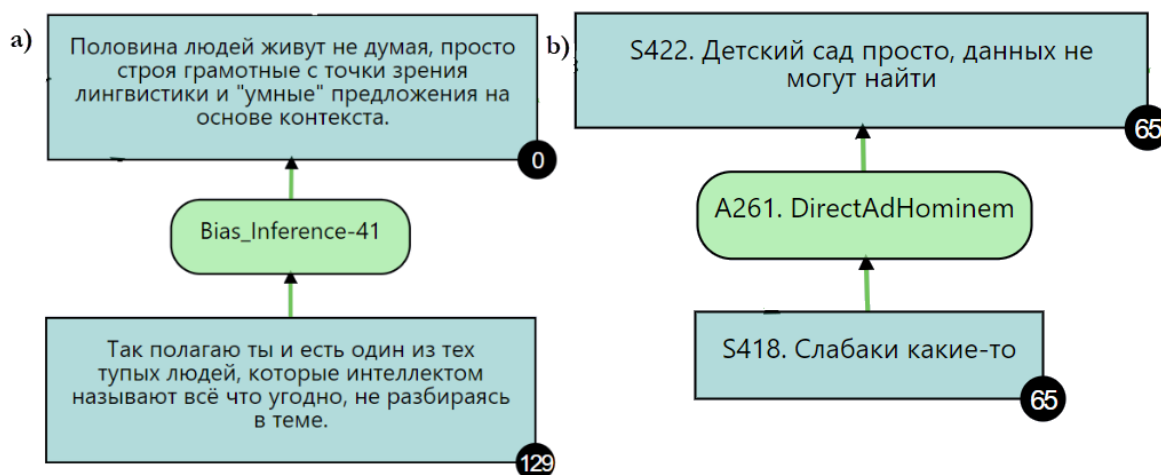


Рис. 5. Не встречающиеся в комментариях рецензентов приёмы

Неупотребительность таких приёмов в научной коммуникации легко объяснить большей формальностью собственно научного функционального стиля, в то время как комментарии на форумах, к которым относится Хабр, позволяют более свободный и сниженный стиль.

4. Обсуждение результатов. Полемика в научном дискурсе возникает в коммуникативных ситуациях, характеризующихся разногласиями относительно приемлемости определенной точки зрения или желательности выполнения определенного действия. Способы ведения полемики весьма разнообразны.

Как и следовало ожидать, полемические приемы существенно зависят от жанра, в рамках которого они используются, что показано в разделе 3 на примере статистики употребительности различных аргументативных приемов. Исключением является логический конфликт, что вполне естественно, поскольку полемика часто связана с отрицанием выдвинутого автором тезиса. С другой стороны, при аннотировании логический конфликт часто используют в случаях, когда тип приёма установить не удается, что сигнализирует о недостаточной систематизации аргументов для данного жанра или тематики.

Существенные отличия заметны в употреблении аргументов, связанных с личностью человека или апелляции к точности формулировок, что подтверждает априорное предположение о том, что в рамках рецензий к научной статье у рецензента меньше поводов полемизировать с употреблением терминов, чем у комментаторов на форуме, поскольку в научной статье употребляются строгие термины, имеющие четкое определение в соответствующей научной области или ясно определенные автором статьи. Также статистически выявляется целый ряд приёмов не употребляемых в жанре рецензий, что подтверждает теоретические предположения.

Полученные в исследовании результаты носят как теоретический характер, способствуя развитию онтологических представлений, исследованию типов рассуждений и обучению аргументации в сфере научного дискурса, так и практический характер, позволяя в дальнейшем применять выявленные особенности конфликтных ситуаций при разметке данных, в вычислительных моделях анализа аргументации, разработке инструкций для генеративных нейросетей, интерпретации результатов анализа.

Заключение. В данной работе предложена предварительная типология и сравнительный анализ полемических аргументативных приёмов на базе разметки текстов двух жанров, относящихся к области научной коммуникации.

Дальнейшие направления исследований будут связаны с выявлением составных аргументативных приёмов. Так, например, критика употребленного термина может сопровождаться выражением собственного мнения рецензента, либо апелляцией к внешнему эксперту.

ному мнению с помощью приёма “От Эксперта”. Едва ли можно ожидать большого количества отсылок к экспертному мнению от участников форума. Другой вариант составного приёма, от которого можно ожидать большей частотности в научном жанре, – это критика, сопровождаемая практическим советом-рекомендацией, т.е. побудительным приёмом “Практический вывод”. Интересно выявить и более длинные статистически релевантные составные цепочки аргументации.

Типология и статистика составных аргументативных приемов позволит выйти на уровень дискурсивных стратегий и рассматривать убедительность аргументации относительно дискурса в целом.

Благодарности. Исследование выполнено за счет гранта Российского научного фонда № 23-11-00261, <https://rscf.ru/project/23-11-00261/>.

Список источников

1. Лисанюк Е.Н. Модели аргументации, рассуждение и убеждение / Е.Н. Лисанюк //РАЦИО.ru, 2016. – № 17(2). – С.35-68.
2. Besnard P., Garcia A., Hunter A. et al. Introduction to structured argumentation. *Argument & computation*, 2014, vol. 5(1), pp.1-4.
3. Toulmin S. *The uses of argument*. Cambridge, Cambridge university press, 2003, 262 p.
4. Walton D. *Argumentation theory: A very short introduction*. Eds. Simari G., Rahwan I. *Argumentation in Artificial Intelligence*, Springer, Boston, 2009, pp. 1-22.
5. Reed C., Walton D. Argumentation schemes in argument-as-process and argument-as-product. *Proc. of Conference celebrating informal logic. OSSA conference archive 75, University of Windsor, Ontario, 2003*.
6. Walton D., Reed C., Macagno F. *Argumentation schemes*. Cambridge, Cambridge University Press, 2008, 456 p.
7. Reed C., Budzynska K., Duthie R. et al. *The Argument Web: an online ecosystem of tools. Systems and services for argumentation. Philosophy & technology*, 2017, vol. 30, pp. 1-24.
8. Peldszus A, Stede M. An annotated corpus of argumentative microtexts. *Argumentation and reasoned action. Proc. of the 1st european conference on argumentation, College Publications, London, 2015, vol. 2, pp. 801-815*.
9. Fishcheva I., Kotelnikov E. Cross-lingual argumentation mining for Russian texts. *Analysis of images, social networks and texts*, 2019, pp. 134-144.
10. Kotelnikov E., Loukachevitch N., Nikishina I. et al. RuArg-2022: Argument mining evaluation. *Computational linguistics and intellectual technologies. Proceedings of the international conference “Dialogue 2022”, Moscow, RSUH, 2022, iss. 21, pp. 333-348*.
11. Pietron M., Olszowski R., Gomułka J. Efficient argument classification with compact language models and ChatGPT-4 refinements. *arXiv preprint*, 2024, available at: <https://arxiv.org/pdf/2403.15473v1> (accessed: 09/10/2024).
12. Musi E., Ghosh D., Muresan S. Towards feasible guidelines for the annotation of argument schemes. In *Proceedings of the 3rd Workshop on argumentation mining, association for computational linguistics, Berlin, 2016, pp. 82–93*.
13. Musi E., Alhindi T., Stede M. et al. A Multi-layer annotated corpus of argumentative text: from argument schemes to discourse relations. *Language resources and evaluation (LREC’2018). Proc. of the 11th international conference, Miyazaki, Japan, 2018, pp. 1629-1636*.
14. Rigotti E., Morasso S.G. Comparing the argumentum model of topics to other contemporary approaches to argument schemes: the procedural and material components. *Argumentation*, 2010, vol. 24(4), pp. 489–512.
15. Feng V.W., Hirst G. Classifying arguments by scheme. In *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies, Portland, Oregon, USA, pp. 987-996*.
16. Park J., Cardie C. A corpus of erulemaking user comments for measuring evaluability of arguments. In *Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018), European Language Resources Association (ELRA), Miyazaki, Japan, 2018, pp. 1623-1628*.
17. Stab Ch., Gurevych I. Parsing argumentation structures in persuasive essays. *Computational linguistics*, 2017, vol. 43(3), pp. 619-660.
18. Potash P., Romanov A., Rumshisky A. Here’s my point: Joint pointer architecture for argument mining. In *Proceedings of the 2017 conference on empirical methods in natural language processing, Copenhagen, Denmark, Association for computational linguistics, 2017, pp. 1364-1373*.
19. Lawrence J. Explainable argument mining. *Ph. D. thesis*, 2021, DOI: 10.15132/20000116.

20. Wagemans J.H.M. Constructing a periodic table of arguments. Ed. Bondy P., Benacquista L. Argumentation, objectivity, and bias. Proceedings of the 11th international conference of the Ontario society for the study of argumentation (OSSA), Windsor, 2016, pp. 1-12.
21. Wagemans J.H.M. How to identify an argument type? On the hermeneutics of argumentative discourse. Journal of pragmatics, 2023, pp. 117-129.
22. Visser J., Lawrence J., Reed C., et al. Annotating argument schemes. Argumentation, 2021, vol. 35, pp. 101-139.
23. Кононенко И.С. Лингвистические аспекты исследования аргументации на основе онтологии / И.С. Кононенко, И.П. Ахмадеева, Е.А. Сидорова // Информационные и математические технологии в науке и управлении, 2020. – № 4 (20). – С. 44-55.
24. Cerutti F., Toniolo A., Norman T.J. et al. AIF-EL – an OWL2-EL compliant AIF ontology. Computational models of argument. Proceedings of COMMA 2018, IOS Press, 2018, vol. 305, pp. 455-456.
25. Сидорова Е.А. Комплексный подход к анализу аргументативных отношений в текстах научной коммуникации / Е.А. Сидорова, И.П. Ахмадеева, Ю.А. Загорюлько и др // Онтология проектирования, 2023. – Т.13. – №4 (50). – С. 562-579.
26. Kononenko I.S., Sery A.S., Shestakov V.K. et al. An approach to classifying Walton's argumentation schemes. 2023 IEEE XVI International scientific and technical conference actual problems of electronic instrument engineering (APEIE), Novosibirsk, 2023, pp. 1540-1545.
27. A Quick start guide to inference anchoring theory (IAT). Centre for argument technology, September 25, 2023, available at: https://www.arg.tech/f/IAT_guidelines_and_tutorials-2023-10.pdf (accessed: 04/20/2024)
28. Тимофеева М.К. Аргументативная разметка корпуса текстов научной интернет-коммуникации: жанровый анализ и исследование типовых моделей рассуждения с помощью платформы ArgNetBank Studio / М.К. Тимофеева, Д.В. Ильина, И.С. Кононенко // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация, 2024. – Т. 22. – № 1. – С. 27-49.

Сидорова Елена Анатольевна. К.ф.-м.н., старший научный сотрудник лаборатории искусственного интеллекта Института систем информатики им. А.П. Ершова СО РАН, доцент кафедры программирования и кафедры систем информатики Новосибирского государственного университета. Основные направления исследований связаны с компьютерной лингвистикой, анализом аргументации, онтологическим инжинирингом и разработкой интеллектуальных систем. Author ID: 146000, SPIN: 3674-4380, ORCID: 0000-0001-8731-3058, lsidorova@iis.nsk.su. 630090, Российская Федерация, г. Новосибирск, проспект Академика Лаврентьева, 6.

Кононенко Ирина Семеновна. Научный сотрудник лаборатории искусственного интеллекта института систем информатики им. А.П. Ершова СО РАН. Основные направления исследований связаны с компьютерной лингвистикой, представлением знаний и анализом аргументации. Author ID: 108210, SPIN: 8680-4501, ORCID: 0000-0001-5057-6807, irina_k@sp.ru, 630090, Российская Федерация, г. Новосибирск, проспект Академика Лаврентьева, 6.

UDC 004.82:81`42

DOI:10.25729/ESI.2024.35.3.002

Ontological analysis of argumentation techniques in scientific discourse

Elena A. Sidorova, Irina S. Kononenko

A.P. Ershov Institute of Informatics Systems SB RAS,
Russia, Novosibirsk, lsidorova@iis.nsk.su

Abstract. The paper is devoted to the analysis of discursive argumentation techniques within the framework of modern trends in the field of reasoning modeling. Reasoning in scientific discourse is organized as a sequence of discursive techniques, i.e. steps of reasoning corresponding to certain mental operations on the object of the subject area. From a formal point of view, the elementary structure of reasoning corresponds to an argument that represents the transition from premises to the thesis being proved. The paper presents an ontological model describing typical arguments – argumentation schemes based on Douglas Walton's theory. The issues of classification of existing modes of reasoning are discussed and an approach to ontological analysis of discourse structures based on

ontological-semantic relations is considered, which provide a conceptual basis for reasoning models and help to understand the nature of entities, relations and reasoning related to a particular context. An annotated corpus of texts that refer to the field of scientific communication was used as material for practical research, including analytical articles with users' comments from the Habr forum (popular science discourse) and scientific articles with comments from reviewers (actually scientific discourse). Of particular interest in this study were the conflicting argumentative techniques used in the texts of the genres under study: Antithesis (`attack` + `on_thesis`) and Counterargument (`attack` + `on_argument`). A statistical analysis and further classification of these techniques depending on the genre were carried out. The obtained classification is based on 12 representative techniques of the first type and 30 techniques of the second type. The ontology, classification of techniques and the corpus are publicly available on the ArgNetBank Studio platform.

Keywords: argumentation analysis, ontology of argumentation, argumentation scheme, polemical techniques

Acknowledgements: This study was supported by the Russian Science Foundation (grant No. 23-11-00261), <https://rscf.ru/project/23-11-00261/>.

References

1. Lisanyuk E.N. Reasoning, persuasion and modelling argumentation [Modeli argumentatsii, rassuzhdenie i ubezhdenie]. *RATIO.ru*, 2016, no. 17(2), pp. 35-68.
2. Besnard P., Garcia A., Hunter A. et al. Introduction to structured argumentation. *Argument & computation*, 2014, vol. 5(1), pp.1-4.
3. Toulmin S. *The uses of argument*. Cambridge, Cambridge university press, 2003, 262 p.
4. Walton D. *Argumentation theory: A very short introduction*. Eds. Simari G., Rahwan I. *Argumentation in Artificial Intelligence*, Springer, Boston, 2009, pp. 1-22.
5. Reed C., Walton D. Argumentation schemes in argument-as-process and argument-as-product. *Proc. of Conference celebrating informal logic. OSSA conference archive 75*, University of Windsor, Ontario, 2003.
6. Walton D., Reed C., Macagno F. *Argumentation schemes*. Cambridge, Cambridge University Press, 2008, 456 p.
7. Reed C., Budzynska K., Duthie R. et al. *The Argument Web: an online ecosystem of tools. Systems and services for argumentation. Philosophy & technology*, 2017, vol. 30, pp. 1-24.
8. Peldszus A, Stede M. An annotated corpus of argumentative microtexts. *Argumentation and reasoned action. Proc. of the 1st european conference on argumentation*, College Publications, London, 2015, vol. 2, pp. 801-815.
9. Fishcheva I., Kotelnikov E. Cross-lingual argumentation mining for Russian texts. *Analysis of images, social networks and texts*, 2019, pp. 134-144.
10. Kotelnikov E., Loukachevitch N., Nikishina I. et al. RuArg-2022: Argument mining evaluation. *Computational linguistics and intellectual technologies. Proceedings of the international conference "Dialogue 2022"*, Moscow, RSUH, 2022, iss. 21, pp. 333-348.
11. Pietron M., Olszowski R., Gomułka J. Efficient argument classification with compact language models and ChatGPT-4 refinements. *arXiv preprint*, 2024, available at: <https://arxiv.org/pdf/2403.15473v1> (accessed: 09/10/2024).
12. Musi E., Ghosh D., Muresan S. Towards feasible guidelines for the annotation of argument schemes. In *Proceedings of the 3rd Workshop on argumentation mining, association for computational linguistics*, Berlin, 2016, pp. 82-93.
13. Musi E., Alhindi T., Stede M. et al. A Multi-layer annotated corpus of argumentative text: from argument schemes to discourse relations. *Language resources and evaluation (LREC'2018)*. *Proc. of the 11th international conference*, Miyazaki, Japan, 2018, pp. 1629-1636.
14. Rigotti E., Morasso S.G. Comparing the argumentum model of topics to other contemporary approaches to argument schemes: the procedural and material components. *Argumentation*, 2010, vol. 24(4), pp. 489-512.
15. Feng V.W., Hirst G. Classifying arguments by scheme. In *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*, Portland, Oregon, USA, pp. 987-996.
16. Park J., Cardie C. A corpus of erulemaking user comments for measuring evaluability of arguments. In *Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018)*, European Language Resources Association (ELRA), Miyazaki, Japan, 2018, pp. 1623-1628.
17. Stab Ch., Gurevych I. Parsing argumentation structures in persuasive essays. *Computational linguistics*, 2017, vol. 43(3), pp. 619-660.
18. Potash P., Romanov A., Rumshisky A. Here's my point: Joint pointer architecture for argument mining. In *Proceedings of the 2017 conference on empirical methods in natural language processing*, Copenhagen, Denmark, Association for computational linguistics, 2017, pp. 1364-1373.
19. Lawrence J. Explainable argument mining. Ph. D. thesis, 2021, DOI: 10.15132/20000116.

20. Wagemans J.H.M. Constructing a periodic table of arguments. Ed. Bondy P., Benacquista L. Argumentation, objectivity, and bias. Proceedings of the 11th international conference of the Ontario society for the study of argumentation (OSSA), Windsor, 2016, pp. 1-12.
21. Wagemans J.H.M. How to identify an argument type? On the hermeneutics of argumentative discourse. Journal of pragmatics, 2023, pp. 117-129.
22. Visser J., Lawrence J., Reed C., et al. Annotating argument schemes. Argumentation, 2021, vol. 35, pp. 101-139.
23. Kononenko I., Akhmadeeva I., Sidorova E. Lingvisticheskie aspekty issledovaniya argumentacii na osnove ontologii [Linguistic aspects of ontology-based argumentation study]. Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii [Information and mathematical technologies in science and management], 2020, no.4 (20), pp. 44-55.
24. Cerutti F., Toniolo A., Norman T.J. et al. AIF-EL – an OWL2-EL compliant AIF ontology. Computational models of argument. Proceedings of COMMA 2018, IOS Press, 2018, vol. 305, pp. 455-456.
25. Sidorova E.A., Akhmadeeva I.R., Zagorulko Yu.A. et al. Kompleksnyj podhod k analizu argumentativnyh otnoshenij v tekstah nauchnoj kommunikacii [An integrated approach to the analysis of argumentative relationships in scientific communication texts]. Ontologiya proyektirovaniya [Ontology of designing], 2023, no.13(4), pp. 562-579.
26. Kononenko I.S., Sery A.S., Shestakov V.K. et al. An approach to classifying Walton's argumentation schemes. 2023 IEEE XVI International scientific and technical conference actual problems of electronic instrument engineering (APEIE), Novosibirsk, 2023, pp. 1540-1545.
27. A Quick start guide to inference anchoring theory (IAT). Centre for argument technology, September 25, 2023, available at: https://www.arg.tech/f/IAT_guidelines_and_tutorials-2023-10.pdf (accessed: 04/20/2024)
28. Timofeeva M.K., Ilina D.V., Kononenko I.S. Argumentativnaja razmetka korpusa tekstov nauchnoj internet-kommunikacii: zhanrovyy analiz i issledovanie tipovyh modelej rassuzhdeniya s pomoshh'ju platformy ArgNetBank Studio [Argumentative annotation of the scientific internet-communication corpus: genre analysis and study of typical reasoning models based on the ArgNetBank Studio platform]. Vestnik NGU. Seriya Lingvistika i mezhkul'turnaya kommunikatsiya [NSU Vestnik. Series linguistics and intercultural communication], 2024, vol. 22, no. 1, pp. 27-49.

Sidorova Elena Anatolievna. PhD, senior researcher of the laboratory of Artificial Intelligence at the A.P. Ershov institute of informatics systems SB RAS (Novosibirsk, Russia), associate professor at Novosibirsk state university. The research areas are related to computational linguistics, intelligent system development, knowledge and ontology engineering. Author ID: 146000, SPIN: 3674-4380, ORCID: 0000-0001-8731-3058, lsidorova@iis.nsk.su, 630090, Russia, Novosibirsk, Acad. Lavrentjev pr., 6.

Kononenko Irina Semenovna. Senior researcher of the laboratory of Artificial Intelligence at the A.P. Ershov institute of informatics systems SB RAS (Novosibirsk, Russia), associate professor at Novosibirsk state university. The research areas are related to computational linguistics, knowledge representation, argument mining. Author ID: 108210, SPIN: 8680-4501, ORCID: 0000-0001-5057-6807, irina_k@cn.ru, 630090, Russia, Novosibirsk, Acad. Lavrentjev pr., 6. Статья поступила в редакцию 05.08.2024; одобрена после рецензирования 11.10.2024; принята к публикации 14.10.2024.

The article was submitted 08/05/2024; approved after reviewing 10/11/2024; accepted for publication 10/14/2024.

УДК 007:159.955:519.72

DOI:10.25729/ESI.2024.35.3.003

Пластичность самоподобных нейронных сетей

Дорогов Александр Юрьевич

Санкт-Петербургский государственный электротехнический университет,

ПАО «Информационные телекоммуникационные технологии»,

Россия, Санкт-Петербург, vaksa2006@yandex.ru

Аннотация. В работе исследуется пластичность многослойных модульных нейронных сетей, обладающих характеристическим свойством самоподобия. Для оценки пластичности используется известное из механики понятие степеней свободы. Число степеней свободы сети оценивается максимальной размерностью операторного многообразия нейронной сети, образованного вариацией параметров нейронных модулей и наличием межмодульных связей. Для получения оценок пластичности нейронные модули рассматриваются как линейные операторы фиксированного ранга. Получены расчетные формулы для вычисления размерности операторного многообразия нейронного модуля вне и в составе сети. Нейронная сеть рассматривается, как двойственный оператор сложной структуры, входом и выходом которого являются векторные пространства. На уровне структурной модели вводится понятие модальных состояний сети, характеризующих размерности векторных подпространств на входе и выходе нейронных модулей в составе сети. Оценка размерности многообразия сети выполняется через ее модальные состояния. Отмечено, что самоподобные сети принадлежат классу слабосвязанных сетей, для которых расчет модальных состояний не вызывает трудностей. Получены точные формулы расчета степени пластичности слабосвязанных нейронных сетей. Результаты анализа использованы для оценки пластичности быстрых нейронных сетей (БНС), и их подмножества - пирамидальных БНС прямой и обратной ориентации.

Ключевые слова: Нейронная сеть, структурная модель, самоподобие, модальные состояния, пластичность, степени свободы

Цитирование: Дорогов А.Ю. Пластичность самоподобных нейронных сетей / А.Ю. Дорогов // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 33-43. – DOI: 10.25729/ESI.2024.35.3.003.

Введение. В нейрофизиологии под пластичностью биологических нейронных сетей понимают способность нейронов сети изменять свои свойства под воздействием внешней среды. В применении к искусственным нейронным сетям это понятие можно трактовать как способность нейронной сети к обучению. В работе [1] автором впервые была предложена оценка пластичности для многослойных нейронных сетей прямого распространения. В основу оценки было положено известное из механики понятие числа степеней свободы.

При анализе пластичности многослойная нейронная сеть рассматривается как некоторая точка-оператор в пространстве операторов. Изменение синаптических весов сети приводит к перемещению этой точки в операторном пространстве. Число возможных степеней свободы операторной точки может служить характеристикой пластичности сети. Данная оценка имеет локальный характер, глобальной оценкой может служить верхняя граница числа степеней свободы по всем допустимым точкам операторного пространства.

Следует отметить, что задача оценки степеней свободы перестраиваемых операторов давно вызывала интерес исследователей, еще в 1971 году Andrews H.C. и Caspari K.L. [2] предложили метод расчета числа степеней свободы для линейных операторов, полученных произведениями слабозаполненных квадратных матриц. Суть метода сводилась к прямому подсчету всех варьируемых коэффициентов, однако, при этом не учитывалось, что при перемножении матриц варьируемые коэффициенты становятся зависимыми, в результате исследователи получили завышенный результат. С позиций сегодняшнего дня матричные произведения, использованные Andrews H.C. и Caspari K.L., можно трактовать, как многослойные нейронные сети с линейными функциями активации и ограниченными связями.

Для нейронных сетей с произвольной структурой в настоящее время неизвестен способ вычисления оценки пластичности, однако для самоподобных нейронных сетей прямого распространения с ядерной организацией [3, 4] такую оценку можно получить для ряда практически важных случаев. Число степеней свободы нейронной сети можно интерпретировать как размерность её операторного многообразия, порождаемого вариацией всех возможных параметров сети. При этом предполагается, что операторное многообразие нейронной сети локально дифференцируемо по всем координатам, т. е. является непрерывным и гладким. Гладкое многообразие локально устроено, как линейное пространство [5], и его размерность оценивается через максимальную размерность локальных пространств. Понятно, что данная оценка не изменится, если вместо полного операторного многообразия нейронной сети рассматривать только его часть (трансверсальную область), в которой размерность локального пространства достигает максимума. Эту выделенную часть обычно называют областью «общего положения».

Концепция «общего положения» в моделировании является выражением фундаментального понятия «типичности» при совместном размещении геометрических образов в векторном пространстве. Операторное многообразие нейронной сети формируется взаимным объединением операторных многообразий нейронных модулей. За счет существующих межмодульных связей эти многообразия имеют множество пересечений. Для оценки пластичности необходимо локализовать трансверсальную область результирующего многообразия и определить ее размерность. Локально каждое операторное многообразие устроено, как некоторое линейное подпространство в пространстве операторов, поэтому многообразия локально пересекаются по векторным подпространствам. Для двух подпространств L_1, L_2 векторного пространства L под общим положением понимается [6] такое их размещение, при котором их пересечение имеет минимальную, а сумма – максимальную размерность (под суммой или объединением подпространств понимается их линейная оболочка). Другой термин для этого же понятия – L_1, L_2 пересекаются трансверсально. Подпространства в «общем положении» либо не пересекаются, а если пересекаются, то размерность их объединения равна:

$$\dim L_1 \cup L_2 = \dim L_1 + \dim L_2 - \dim L_1 \cap L_2. \quad (1)$$

Название «общее положение» обусловлено тем, что в некотором смысле большинство пар подпространств L_1, L_2 находятся в общем положении, а другие расположения являются вырожденными.

Поскольку размерность многообразия оценивается через локальную размерность линейного пространства, то в контексте оценки пластичности можно ограничиться линейными операторами. Для модульной нейронной сети задача расчёта степени пластичности разделяется на две подзадачи: в первой требуется определить пластичность отдельных модулей, а во второй – используя полученные данные и информацию о структуре сети, определить пластичность всей сети, следуя при этом принципу трансверсальности при объединении операторных многообразий. Рассмотрим обе эти задачи последовательно.

2. Операторные многообразия нейронных модулей. Будем полагать, что в контексте оценки пластичности нейронной сети нейронный модуль описывается линейным оператором. Пусть E и D – ассоциированные пространства входа и выхода для модуля A с размерностями p и g по входу и выходу. Ограничиваясь моделью общего положения, будем считать, что при варьировании параметров операторное многообразие модуля совпадает с многообразием линейных отображений $A_r: E \rightarrow D$ ранга r .

Как известно [7], полное множество линейных операторов, действующих из пространства E в D , изоморфно тензорному произведению $E \otimes D$, которое является линейным пространством размерности $p \cdot g$. Согласно теореме о структуре линейного отображения [6] для каждого оператора ранга r существуют такие прямые разложения пространств:

$$E = E_0 \oplus E_1 \Leftrightarrow D_1 \oplus D_0 = D, \tag{2}$$

что D_1 изоморфно E_1 и имеет размерность r , а E_0 составляет ядро отображения. Подпространство D_0 называют коядром, оно характеризует степень неопределённости (неоднозначности) оператора.

Из разложения (2) нетрудно видеть, что подмножество операторов A_r представляет собой объединение подмножества операторов ранга r , осуществляющих отображение из пространства E в D_1 , и подмножества операторов ранга r , осуществляющих отображение из E_1 в D . На языке тензорных произведений это можно записать так:

$$A_r \cong (E \otimes D_1) \cup (E_1 \otimes D).$$

Отсюда, следуя правилу вычисления размерности объединения подпространств общего положения (1), получим:

$$\dim A_r = \dim(E \otimes D_1) + \dim(E_1 \otimes D) - \dim(E \otimes D_1) \cap (E_1 \otimes D). \tag{3}$$

Пересечением множеств является подмножество операторов, осуществляющих отображение из подпространства E_1 в D_1 , поэтому

$$\dim(E \otimes D_1) \cap (E_1 \otimes D) = r^2.$$

Поскольку $\dim E = p$, $\dim D = g$, $\dim E_1 = \dim D_1 = r$, то из (3) следует:

$$\dim A_r = pr + gr - r^2. \tag{4}$$

В частности, для операторного многообразия полного ранга, когда $r = \min(p, g)$, из данного выражения получим очевидное значение $\dim A_r = p \cdot g$. Аналогичные соотношения для определения размерности многообразия матриц ранга r , представлены в работе [8].

3. Модуль двойственного функционирования в составе сети. Обозначим $A_r^\# : D \rightarrow E$ множество всех операторов обратного распространения ранга r , действующих из пространства D в пространство E . Разложение (2) симметрично для класса прямых и обратных отображений, поэтому также представляет собой многообразие и существует естественный изоморфизм $A_r \cong A_r^\#$, который задаётся совпадением пар (E_1, D_1) . Данный изоморфизм является выражением двойственности в представлении модуля (см. рис. 1).

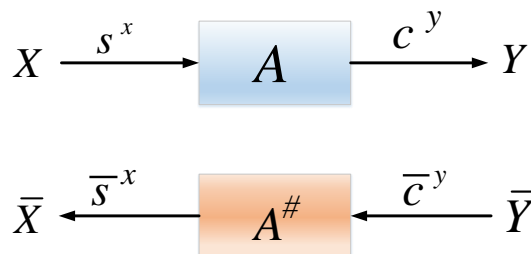


Рис. 1. Нейронные модули в прямой и двойственной сети

Возможность тензорного представления множества операторов позволяет в расчетах перейти от операторного многообразия к одному оператору фиксированного ранга, отображающего входное пространство векторов в выходное векторное пространство. Зная размерности этих пространств для прямого и двойственного операторов, размерность многообразия операторов можно вычислить, следуя выражению (4). Условия общего положения теперь предъявляются к корректному выбору оператора.

Таким образом, будем полагать, что операторный модуль общего положения ранга r , размерностью p и g по входу и выходу находится под воздействием пространств E и D . В режимах прямого и обратного функционирования размерность выходного пространства операторного модуля общего положения не может быть больше r . Следовательно, модуль может реализовать не более, чем pr степеней свободы при прямом функционировании и не более, чем gr – при обратном. При этом r^2 степеней свободы являются общими. Дуальный переход от многообразия операторов к векторным пространствам, ассоциированным с каждым модулем сети, естественным образом вводит сигнальные состояния на структурной модели нейронной сети. Размерность сигнальных подпространств в различных сечениях двойственной сети теперь будем рассматривать как модальные состояния структурного уровня (от латинского *modus* – размер). Операторные модули общего положения осуществляют преобразование входного модального состояния в выходное в прямой и двойственной сетях.

Если модуль находится в составе сети, то размерности входных сигнальных пространств $X \subseteq E$ и $\bar{Y} \subseteq D$ в прямой и двойственной сетях в общем случае меньше максимально возможных (т.е. p и g). Это ограничивает возможности по реализации степеней свободы модуля. Действующее число степеней свободы модуля в составе сети по аналогии с (4) будет равно

$$\dim \hat{A}_r = s^x \hat{r} + \bar{c}^y \hat{r} - \hat{r}^2, \quad (5)$$

где s^x, \bar{c}^y – размерности входных подпространств в прямой и двойственной сетях, $\hat{r} = \min(r, s^x, c^y)$ – действующий ранг нейронного модуля.

При системном анализе наибольший интерес представляют потенциально достижимые возможности нейронной сети. Понятно, что с увеличением операторных рангов модулей такие возможности увеличиваются, поэтому при анализе можно полагать, что всегда выполняется равенство $r = \min(p, g)$. В дифференциальной геометрии такой оператор называется регулярным или правильным; этот же термин будем использовать для нейронного модуля. Для регулярного модуля в составе сети выполняется условие $r \geq \min(s^x, c^y)$ и, следовательно, $\hat{r} = \min(s^x, \bar{c}^y)$, тогда из (5) нетрудно получить, что

$$\dim \hat{A}_r = s^x \bar{c}^y. \quad (6)$$

Для операции нахождения минимума в дальнейшем будем использовать обозначение « \circ », иначе называемое логическим умножением. Например, с этим обозначением условие регулярности модуля будет записываться в виде $r = p \circ g$.

4. Пластичность модульных нейронных сетей с биективными связями.

Будем полагать, что многообразие операторов в модульной нейронной сети образуется только за счёт вариации параметров модулей. Межмодульные связи при этом считаются фиксированными.

4.1. Влияние модулей. Выделим в сети некоторый модуль A ранга r и предположим, что его параметры варьируются, в то время как параметры всех остальных модулей фиксированы в общем положении. Поскольку модуль находится в составе сети, то размерности пространств состояний модуля в прямой и двойственной сетях в общем случае меньше размерностей модуля по входу и выходу. Действующее число степеней свободы модуля обозначим $S(A)$. Выделяя подобным образом и поодиночно варьируя параметры остальных модулей, получим, что вклад всех модулей в общее число степеней свободы сети определяется суммой:

$$\sum_A S(A).$$

4.2. Влияние связей. Биективные межмодульные связи устанавливают точное и однозначное соответствие между модальными состояниями смежных модулей, т. е. каждая связь, рассматриваемая как двойственное отображение, является регулярной и её влияние можно оценивать выражением (6). Однако, в отличие от модулей, связи фиксированы, и поэтому их присутствие в сети приводит к «связыванию» степеней свободы смежных нейронных модулей. Из этого следует, что каждая фиксированная межмодульная связь (допустим, между модулями $A \Rightarrow B$) ранга r^{AB} уменьшает общее количество степеней свободы сети на величину

$$s^B \bar{c}^A = (c^A \circ r^{AB})(\bar{s}^B \circ r^{AB}),$$

где s^B, c^A и \bar{c}^A, \bar{s}^B размерности пространств модальных состояний в прямой и двойственной сетях (вторая форма записи более удобна при конусе связей). Рис. 2 иллюстрирует действие фиксированной межмодульной связи в сети двойственного распространения.

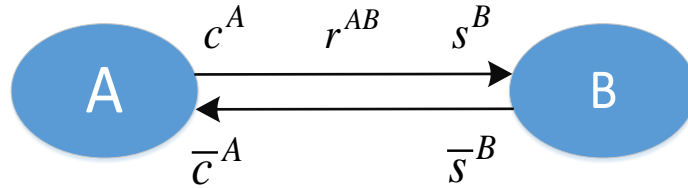


Рис. 2. Межмодульная связь в нейронной сети двойственного распространения

Таким образом, формула расчёта числа степеней свободы для всей модульной сети будет иметь вид:

$$S(H) = \sum_A S(A) - \sum_{A \Rightarrow B} \sum (c^A \circ r^{AB})(\bar{s}^B \circ r^{AB}), \tag{7}$$

а если все модули регулярны, то:

$$S(H) = \sum_A s^A \bar{c}^A - \sum_{A \Rightarrow B} \sum (c^A \circ r^{AB})(\bar{s}^B \circ r^{AB}). \tag{8}$$

В последних выражениях двойная сумма распространяется на все существующие связи.

5. Расчёт модальных состояний в самоподобных нейронных сетях.

Выражения (7), (8) можно использовать для расчёта степени пластичности только в том случае, если известны модальные состояния нейронной сети. Модальные состояния в прямой и двойственной сетях порождаются наличием на их входах сигнальных пространств. Для оценки потенциальных свойств нейронной сети будем полагать, что эти пространства имеют максимально возможную размерность, определяемую размерностью терминальных полей сети.

$$s^{inp} = N, \quad \bar{c}^{out} = M.$$

Здесь N – размерность нейронной сети по входу, M – размерность нейронной сети по выходу. Напомним, что выход прямой сети является одновременно входом двойственной сети, а состояния s^{inp}, \bar{c}^{out} являются источниками, воздействующими на прямую и двойственную сеть. Схема формирования модальных состояний двойственной сети показана на рис. 3.

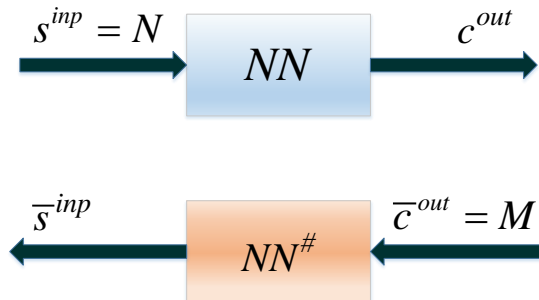


Рис. 3. Схема формирования модальных состояний в прямой и двойственной нейронной сети

Рассмотрим конус связей к некоторой вершине B в модульной сети прямого распространения (см. рис. 4). Пусть эта вершина соответствует оператору ранга r^B с размерностью p^B по входу и g^B по выходу (рис. 5).

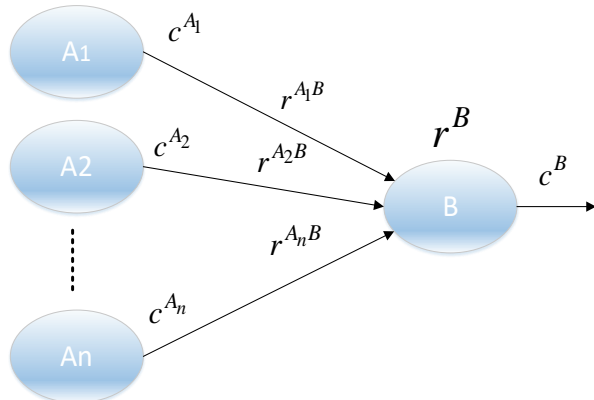


Рис. 4. Конус связей в сети прямого распространения

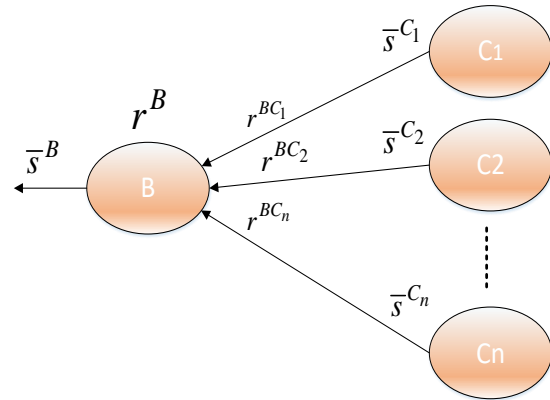


Рис. 5. Конус связей в сети обратного распространения

Модальные состояния для конуса связей в общем виде определяются выражениями:

$$c^B = s^B \circ r^B, \quad s^B = \bigcup_{A_i \in \Gamma^{-1}(B)} c^{A_i} \circ r^{A_i B},$$

где символ \bigcup условно показывает, что размерность s^B следует определять для объединения подпространств, порождаемых модулями окрестности $\Gamma^{-1}(B)$. Размерности этих подпространств известны и равны $c^{A_i} \circ r^{A_i B}$, однако размерность их объединения в общем случае вычислить не удаётся, поскольку в произвольной структуре пространства могут иметь пересечения, обусловленные наличием предшествующих связей между модулями.

В случае самоподобных сетей задача существенно упрощается, поскольку по теореме «О параллельных путях» [3] для любого модуля самоподобной сети существует единственный путь, связывающий его с терминальной вершиной. Этот путь выделяет в терминальном пространстве независимое подпространство. Размерность этого подпространства легко вычислить: оно будет равно логическому произведению рангов всех дуг пути от терминальной вершины до данного модуля. Поскольку все выделенные таким образом подпространства не пересекаются, то объединение подпространств перед модулем B фактически будет прямой суммой, а это означает, что формула для вычисления размерности трансформируется к виду:

$$c^B = s^B \circ r^B, \quad s^B = \sum_{A_i \in \Gamma^{-1}(B)} c^{A_i} \circ r^{A_i B} = p^B. \quad (9)$$

Поскольку модули регулярны, то $c^{A_i} \circ r^{A_i B} = r^{A_i B}$. Более того, в самоподобных сетях все связи инъективны, поэтому сумма (9) будет равняться входной размерности модуля, т.е. $s^B = p^B$. Следуя принципу двойственности, можно записать аналогичные выражения для расчёта модальностей для сети обратного распространения (см. рис.5):

$$\bar{s}^B = \bar{c}^B \circ r^B, \quad \bar{c}^B = \sum_{C_i \in \Gamma(B)} \bar{s}^{C_i} \circ r^{C_i B} = g^B, \quad (10)$$

где g^B - размерность модуля по выходу.

5.1. Пластичность быстрых нейронных сетей. Быстрое перестраиваемое преобразование [3] будем рассматривать как модульную быструю нейронную сеть (БНС), где роль модулей выполняют базовые операции (нейронные ядра). Эта сеть является самоподобной. По построению быстрого преобразования ранги всех межмодульных связей равны единице, а все ядра в пределах слоя имеют совпадающие структурные характеристики. Обозначим через

(p_m, g_m) размерности ядер слоя m , а через k_m – их число. Нетрудно заметить, что при единичных рангах каждая фиксированная связь удаляет одну степень свободы. Учитывая это обстоятельство и используя (9), (10) из формулы (8), получим:

$$S(H) = \sum_{m=0}^{n-1} p_m g_m k_m - \sum_{m=0}^{n-2} D_m, \quad (11)$$

где D_m – количество связей в межслойном переходе с номером m , n – число нейронных слоев. Например, для сети, топологическая и структурная модели которой показаны на рис. 6 и рис.7, число степеней свободы равно $S(H) = (3 \cdot 2 \cdot 4 + 2 \cdot 2 \cdot 4 + 2 \cdot 2 \cdot 4) - (8 + 8) = 40$.

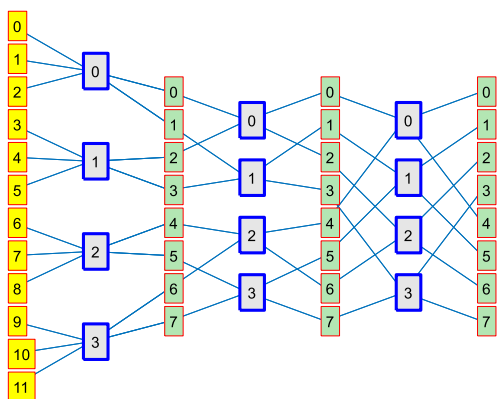


Рис. 6. Топологическая модель трехслойной БНС

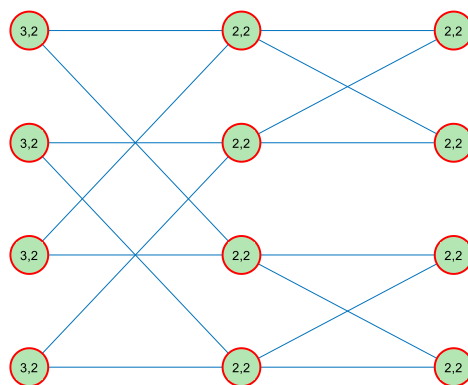


Рис. 7. Структурная модель трехслойной БНС

5.2. Пластичность самоподобной пирамидальной нейронной сети. Ключевая идея самоподобных пирамидальных нейронных сетей [9] заключается в изменении топологии БНС таким образом, чтобы выделить сетевые компоненты, информационная ёмкость которых используется полностью и отсечь компоненты, информационная ёмкость которых не используется при обучении, сохранив при этом условия самоподобия. В результате нейронная сеть упрощается, быстродействие её увеличивается, и глубина обучения становится максимально-возможной. На рис. 8 и рис. 9 показаны топологическая и структурная модели самоподобной пирамидальной нейронной сети. На структурной модели ранги всех связей равны 1.

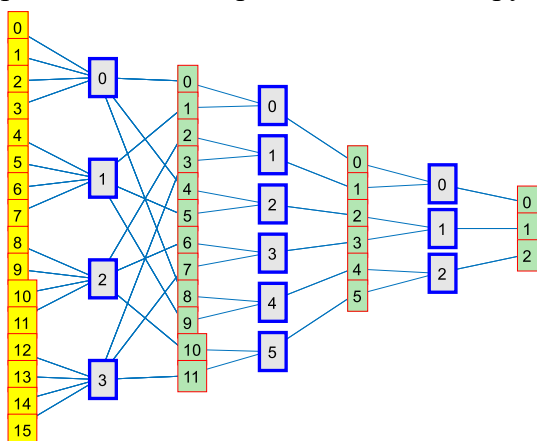


Рис. 8. Топологическая модель пирамидальной самоподобной сети

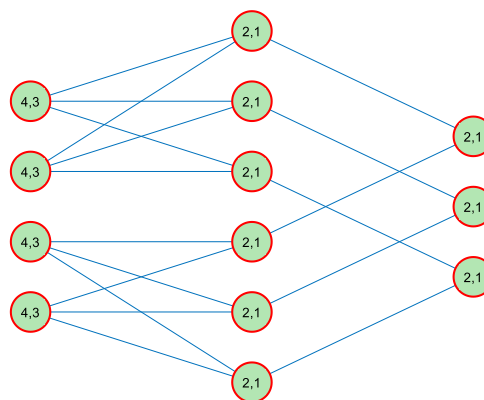


Рис. 9. Структурная модель пирамидальной самоподобной сети

Данная сеть может обучиться к трём эталонным функциям, заданным на интервале длиной 16. Используя расчетную формулу (11), непосредственно из структурной модели получим

$S(H) = 4 \cdot 3 \cdot 4 + 2 \cdot 1 \cdot 6 + 2 \cdot 1 \cdot 3 - 12 - 6 = 48$. При произвольном выборе трёх функций необходимо задать 48 значений, что совпадает с полученным числом степеней свободы пирамидальной сети. Таким образом, рассматриваемая сеть является глубокой в том смысле, что её потенциал обучения используется полностью, и покрывает все допустимое многообразие эталонных сигналов.

5.3. Информационная емкость нейросетевой памяти образов. В [9] показано, что самоподобная пирамидальная сеть обратной ориентации может быть эффективно использована для хранения одномерных и многомерных образов. Пластичность нейронной сети здесь трактуется как информационная емкость памяти. На рис. 10 и рис. 11 приведен пример нейросетевой памяти в топологическом и структурном представлении пирамидальной сети. Ранги всех связей структурной модели равны 1.

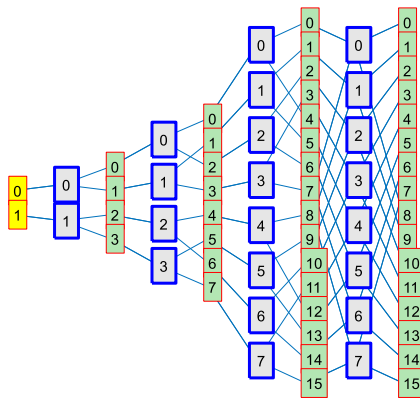


Рис. 10. Топологическая модель нейросетевой памяти образов

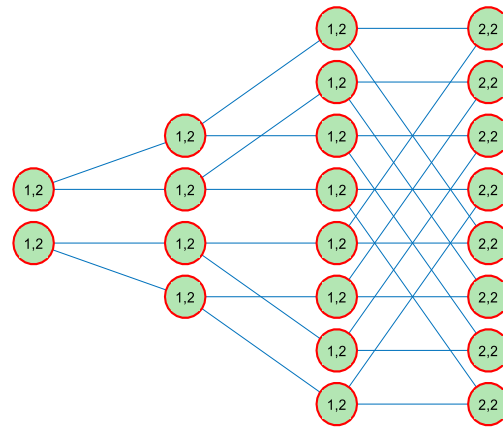


Рис. 11. Структурная модель нейросетевой памяти образов

Данная нейронная сеть обеспечивает хранение двух произвольных дискретных функций, заданных на интервале длиной $M = 16$. Используя расчетную формулу (11), непосредственно из структурной модели получим $S(H) = (1 \cdot 2) \cdot 14 + (2 \cdot 2) \cdot 8 - 28 = 32$. Для задания двух функций на интервале длиной 16 требуется определить 32 значения. Таким образом, сеть полностью использует свой потенциал для хранения функций и поэтому относится к категории сетей с глубокой степенью обучения. В отличие от кристаллов цифровой памяти с последовательным хранением данных, где восстановление образа происходит за счёт последовательного опроса ячеек, в памяти на БНС все пиксели образа восстанавливаются одновременно, что потенциально обеспечивает сверхвысокое быстродействие при считывании памяти.

Заключение. Пластичность нейронных сетей характеризует дифференциальную способность нейронной сети к обучению. Можно сказать, что качество одной нейронной сети лучше другой, если первая при прочих равных условиях обладает лучшей пластичностью. Но важно заметить, что дифференциальные оценки дают только линейное приближение для локальной области операторного пространства нейронной сети. Нелинейные свойства нейронной сети при этом игнорируются. Можно считать, что полученные численные значения являются нижней границей для оценки информационной емкости сети. В настоящее время не существует методов, позволяющих корректно учесть вклад нелинейностей в интеллектуальный потенциал нейронных сетей произвольной архитектуры. Однако наблюдается тенденция упрощения нелинейных элементов, например, замена их полулинейными элементами типа ReLU [10], что во многом обусловлено технологическими ограничениями при аппаратной реализации нейронных сетей. Элементы ReLU в этом случае выполняют роль управляемых ключей, а нейронная сеть работает в кусочно-линейном режиме.

В данной работе получены формулы расчета степени пластичности самоподобных многослойных нейронных сетей. Приведенные выражения позволяют определить точное значение степени пластичности на основе структурных характеристик нейронной сети. Расчетные формулы базируются на теореме «О параллельных путях», которая утверждает, что структурные модели самоподобных сетей не содержат параллельных путей. Сети, в которых отсутствуют параллельные пути, называются также слабосвязанными [11, 12]. Класс слабосвязанных сетей несколько шире класса самоподобных сетей, в частности, к этому классу относятся полностью связанные многослойные сети, однослойные персептроны, двухслойные ядерные сети с произвольной структурой. Полученные расчетные формулы применимы ко всему классу регулярных слабосвязанных сетей. В [11] представлен также графический расчет числа степеней свободы для слабосвязанных сетей общего вида.

Список источников

1. Дорогов А.Ю. Пластичность многослойных нейронных сетей / А.Ю. Дорогов // Известия вузов. Приборостроение, 1998. – Т.41. – №4. – С.36-41.
2. Andrews H.C., Caspari K.L. Degrees of freedom and modular structure in matrix multiplication. IEEE. Tr. Compt, 1971, vol. C-20, pp.113-141.
3. 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.
4. Дорогов А.Ю. Быстрые преобразования и самоподобные нейронные сети глубокого обучения. Часть 2. Методы обучения быстрых нейронных сетей/ А.Ю. Дорогов // Информационные и математические технологии в науке и управлении, 2024.–№ 1(33). –С.5-19. –DOI:10.25729/ESI.2024.33.1.001.
5. Фоменко А.Т. Наглядная геометрия и топология: Математические образы в реальном мире – 2-е изд. / А.Т. Фоменко. – М.: Изд-во Моск.ун-та, Изд-во "ЧеРо", 1998. – 416 с.
6. Кострикин А.И. Линейная алгебра и геометрия: учеб. пособ. для вузов. 2-е изд. перераб. / А.И. Кострикин, Ю.И. Манин. – М.: Наука, 1986. – 304 с.
7. Ефимов Н.В. Линейная алгебра и многомерная геометрия / Н.В. Ефимов, Э.Р. Розендорн. – М.: Наука, 1970. – 528 с.
8. Дубровин Б.А. Современная геометрия: методы и приложения. Т. 2: Геометрия и топология многообразий. 5-е изд., испр. / Б.А. Дубровин, С.П. Новиков, А.Т. Фоменко. – М.: Эдиториал УРСС, 2001. – 296 с.
9. Дорогов А.Ю. Быстрые преобразования и самоподобные нейронные сети глубокого обучения. Часть 3. Пирамидальные нейронные сети с глубокой степенью обучения / А.Ю. Дорогов // Информационные и математические технологии в науке и управлении, 2024. – № 2(34). – С. 19-32. – DOI:10.25729/ESI.2024.34.2.002.
10. Liu Danqing A practical guide to relu. Medium, available at: <https://readmedium.com/a-practical-guide-to-relu-b83ca804f1f7> (accessed: 04/23/2024).
11. Дорогов А.Ю. Теория и проектирование быстрых перестраиваемых преобразований и слабосвязанных нейронных сетей. / А.Ю. Дорогов. – СПб.: Политехника, 2014. – 328 с.
12. Дорогов А. Ю. Самоподобные нейронные сети быстрого обучения/ А.Ю. Дорогов. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2024. – 188 с.

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.

Plasticity of self-similar neural networks

Alexander Yu. Dorogov

St. Petersburg State Electrotechnical University,
PJSC “Information Telecommunication Technologies” (“Inteltech”),
Russia, St. Petersburg, vaksa2006@yandex.ru

Abstract. The plasticity of multilayer modular neural networks with the characteristic property of self-similarity is investigated in the paper. The concept of degrees of freedom, known from mechanics, is used to assess plasticity. The number of degrees of freedom of the network is estimated by the maximum dimension of the operator manifold of the neural network formed by variation of the parameters of neural modules and the presence of intermodule connections. To obtain plasticity estimates, neural modules are considered as linear operators of fixed rank. Calculation formulas for calculating the dimension of the operator manifold of a neural module outside and within the network are obtained. A neural network is considered as a dual operator of a complex structure, the input and output of which are vector spaces. At the level of the structural model, the concept of modal network states is introduced, characterizing the dimensions of vector subspaces at the input and output of neural modules in the network. The dimensionality of the network manifold is estimated through its modal states. It is noted that self-similar networks belong to a class of weakly coupled networks for which the calculation of modal states does not cause difficulties. Exact formulas for calculating the degree of plasticity of weakly coupled neural networks are obtained, the results of the analysis are used to assess the plasticity of fast neural networks (BNS), and their subsets - pyramidal BNS of direct and reverse orientation.

Keywords: Neural network, structural model, self-similarity, modal states, plasticity, degrees of freedom

References

1. Dorogov A.Yu. Plastichnost' mnogosloynnykh neyronnykh setey [Plasticity of multilayer neural networks]. *Izvestiya vuzov. Priborostroyeniye* [Journal of Instrument Engineering], 1998, vol 41, no. 4, pp. 36-41.
2. Andrews H.C., Caspari K.L. Degrees of freedom and modular structure in matrix multiplication. *IEEE. Tp. Compt*, 1971, vol. C-20, pp.113-141.
3. Dorogov A.Yu. Bystrye preobrazovaniya i samopodobnye neyronnye seti glubokogo obucheniya. Chast' 1. Stratifitsirovannye modeli samopodobnykh neyronnykh setey i bystrykh preobrazovaniy [Fast transformations and self-similar deep learning neural networks. Part 1. Stratified models of self-similar neural networks and fast transformations]. *Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii* [Information and mathematical technologies in science and management], 2023, no. 4(32), pp.5-20, DOI:10.25729/ESI.2023.32.4.001.
4. Dorogov A.Yu. Bystrye preobrazovaniya i samopodobnye neyronnye seti glubokogo obucheniya. Chast' 2. Metody obucheniya bystrykh neyronnykh setey [Fast transformations and self-similar neural networks of deep learning. Part 2. Methods of training fast neural networks]. *Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii* [Information and mathematical technologies in science and management], 2024, no. 1(33), pp. 5-19, DOI:10.25729/ESI.2024.33.1.001.
5. Fomenko A.T. Naglyadnaya geometriya i topologiya: Matematicheskie obrazy v real'nom mire. [Mathematical images in the real world]. 2nd ed. M., Izdatel'stvo Moskovskogo universiteta, Izdatel'stvo CheRo [Publ. Moscow university CheRo], 1998, 416 pp.
6. Kostrikin A.I., Manin Yu.I. Lineynaya algebra i geometriya: ucheb. posob. dlya vuzov [Linear algebra and geometry. Studies. help. for universities]. 2nd ed. Moscow, Nauka [Science], 1986, 304 pp.
7. Efimov N.V., Rozendorn E. R. Lineynaya algebra i mnogomernaya geometriya [Linear algebra and multidimensional geometry]. Moscow, Nauka [Science], 1970, 528 pp.
8. Dubrovin B.A., Novikov S.P., Fomenko A.T. Sovremennaya geometriya: metody i prilozheniya. T. 2: Geometriya i topologiya mnogoobraziy. [Modern geometry: methods and applications. Vol. 2: Geometry and topology of manifolds]. 5nd ed. Moscow, Editorial URSS, 2001, 296 pp.
9. Dorogov A.Yu. Bystrye preobrazovaniya i samopodobnye neyronnye seti glubokogo obucheniya Chast' 3. Piramidal'nye neyronnye seti s glubokoy stepen'yu obucheniya [Fast transformations and self-similar deep learning neural networks Part 3. Pyramid neural networks with a deep learning degree]. *Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii* [Information and mathematical technologies in science and management], 2024, no. 2(34), no. 1(33), pp. 19-32, DOI:10.25729/ESI.2024.34.2.002.
10. Liu Danqing A practical guide to relu. Medium, available at: <https://readmedium.com/a-practical-guide-to-relu-b83ca804f1f7> (accessed: 04/23/2024).

11. Dorogov A.Yu. Teorija i proektirovanie bystryh perestraivaemyh preobrazovanij i slabosvjazannyh nejronnyh setej [Theory and design of fast tunable transformations and loosely coupled neural networks]. Spb. Publ Politehnika. 2014, 328pp.
12. Dorogov A.Yu. Samopodobnye nejronnye seti bystrogo obucheniya [Self-similar neural networks of fast learning]. SPb., Izd-vo SPbGETU «LETI» [St. Petersburg, Publishing house of SPbSETU “LETI”], 2024, 188 p.

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 12.05.2024; одобрена после рецензирования 23.07.2024; принята к публикации 30.09.2024.

The article was submitted 05/12/2024; approved after reviewing 07/23/2024; accepted for publication 09/30/2024.

УДК 004.93.12

DOI:10.25729/ESI.2024.35.3.004

Анализ снимков кожных новообразований с применением комбинированной архитектуры сверточных нейронных сетей

Милантьев Сергей Андреевич^{1,2}, Святкина Виталия Игоревна², Бессмертный Игорь Александрович¹, Зайченко Кирилл Вадимович²

¹Университет ИТМО, Россия, Санкт-Петербург, nevdev@niuitmo.ru

²ИАП РАН, Россия, Санкт-Петербург

Аннотация. В данном исследовании рассматривается возможность применения комбинированных архитектур сверточных нейронных сетей для анализа кожных новообразований. Разработаны архитектуры моделей для извлечения дополнительных признаков, связанных с характером формы кожных новообразований. Проведена оптимизация моделей, а также архитектуры, с целью минимизации ошибок первого и второго рода для редких кожных новообразований. В процессе обучения использовались наборы данных ISIC2017-2020, MED-NODE, SD-198, 7-Point Criteria Database, Light Field Image Dataset of Skin Lesions, PH², датасет ИАП РАН. Для обучения классификационных моделей использовались оптимизатор AdamW, функции потерь FocalLoss и scheduler CosineAnnealingWarmRestarts. Для обучения сегментационных моделей применялась функция потерь BCEDice. Оценка моделей проводилась с использованием взвешенных классификационных метрик, таких, как Recall, Precision и F1-score. В качестве рассматриваемой концепции моделей учитывалась их устойчивость на этапе валидации. Модели, использующие дополнительные сверточные нейронные сети для извлечения признаков формы новообразования, демонстрируют более высокие показатели метрик, а также имеют меньшую сумму ошибок первого и второго рода для редких заболеваний по сравнению с обычными классификационными моделями. Полученные результаты могут найти применение в анализе медицинских задач с дисбалансом данных в обучающем наборе данных.

Ключевые слова: кожные новообразования, сверточные нейронные сети, анализ кожных поражений, дисбаланс классов, многоспектральная обработка изображений

Цитирование: Милантьев С.А. Анализ снимков кожных новообразований с применением комбинированной архитектуры сверточных нейронных сетей / С.А. Милантьев, В.И. Святкина, И.А. Бессмертный, К.В. Зайченко // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 44-56. – DOI: 10.25729/ESI.2024.35.3.004.

Введение. Ранняя диагностика злокачественных новообразований кожи является актуальным исследовательским направлением в области дерматологии. Особое внимание уделяется классификации редких кожных поражений в задаче многоклассовой классификации. Анализ редких кожных новообразований представляет собой важное направление исследований, поскольку эти патологии являются малоизученными и вызывают особый интерес у исследователей. В настоящее время предпринимаются значительные усилия для развития новых методов и технологий с целью повышения точности распознавания и классификации редких кожных новообразований. В частности, применение методов искусственного интеллекта становится все более значимым при разработке инновационных подходов к диагностированию, что позволяет улучшить качество и надежность получаемой информации, а также обеспечить высокую точность в классификации редких кожных новообразований.

Редкие заболевания с точки зрения применения методов искусственного интеллекта представляют собой малочисленные классы в наборе данных, что приводит к значительному дисбалансу в распределении классов. Дисбаланс данных является естественной особенностью многих наборов данных и приводит к низким значениям метрик качества модели для редких классов. Тем не менее, существует ряд подходов для снижения ошибок при дисбалансе классов, хотя даже их применение не обеспечивает идеальную модель для классификации разнообразных изображений, таких, как медицинские снимки кожных новообразований. В связи с

этим, основной целью данного исследования является применение сверточных нейронных сетей для улучшения показателей метрик классификации редких кожных новообразований.

1. Методы и подходы. С каждым годом методы машинного и глубокого обучения играют все более значимую роль в классификации кожных поражений, обеспечивая точность и скорость автоматизированной диагностики. На сегодняшний день для классификации новообразований кожи известны следующие подходы:

1. Классические методы извлечения и классификации признаков. Среди ранних подходов выделяются методы получения гистограмм ориентированных градиентов (HOG), выявления локальных бинарных паттернов (LBP) и анализа цветовых гистограмм новообразований. Эти признаки классифицировались с помощью алгоритмов, таких, как метод опорных векторов (SVM), алгоритм ближайших соседей (KNN) и многослойные перцептроны (MLP). Несмотря на устаревание данных подходов и их ограниченную точность по сравнению с современными методами, они все еще встречаются в ряде современных исследований для решения специфических задач.
2. Применение сверточных нейронных сетей и трансферного обучения. С 2017 года, после появления крупных публичных соревнований, таких, как ISIC на платформе Kaggle, сверточные сети (CNN) с трансферным обучением стали доминировать в области анализа кожных новообразований. Эти сети, предобученные на больших наборах данных (например, ImageNet или других медицинских датасетах), показали высокие показатели метрик при решении медицинских задач с дообучением на специализированных наборах данных изображений кожных поражений.
3. Мультимодальные методы. Мультимодальные модели, которые объединяют различные типы данных, такие, как изображения, метаданные о пациенте и текстовые описания диагнозов, обеспечивают значительное улучшение качества классификации. Разработка мультимодальных подходов продолжается, и они часто демонстрируют лучшие показатели метрик в сравнении с классическими сверточными нейронными сетями. В мультимодальный подход входят и модели из данного исследования.
4. Методы на основе Vision Transformers (ViT). Введение Vision Transformers внесло значительный вклад в анализ изображений, так как они способны выявлять глобальные закономерности, анализируя зависимость между всеми частями изображения одновременно, а не только в пределах рецептивного поля, как это делают CNN. Благодаря этому, ViT более эффективно выявляют сложные паттерны, что особенно важно для диагностики кожных заболеваний.
5. Использование генеративных моделей. Генеративные модели GAN (Generative Adversarial Networks) и диффузионные модели находят применение для создания синтетических данных, особенно при ограниченности доступных изображений редких классов кожных новообразований. Генерация качественных синтетических изображений позволяет улучшить качество обучения моделей за счет увеличения объема данных и уменьшения дисбаланса классов.

Начальной точкой данного исследования является анализ данных, полученных с использованием метода мультиспектральной обработки изображений. Метод мультиспектральной обработки изображений предусматривает использование последовательностей изображений, где каждый элемент освещается группами диодов разных длин волн. Для получения таких последовательностей используется специальное устройство, разработанное в ИАП РАН под руководством Зайченко Кирилла Вадимовича.

Метод мультиспектральной обработки позволяет извлекать пространственную информацию о характере новообразования. Для этого осуществляется получение последовательности

монохромных изображений, на каждом из которых представлено распределение интенсивности светового излучения на определенной длине волны.

Таким образом, применение нескольких групп освещения с различными длинами волн позволяет получить дополнительную спектральную информацию, которая дополняет уже имеющуюся пространственную информацию.

Предполагается, что форма поражения является важным фактором, определяющим различия при классификации кожных новообразований. В данном исследовании не использовались данные, полученные с использованием метода мультиспектральной обработки изображений, поэтому было решено провести ряд экспериментов с целью улучшения показателей метрик в задаче многоклассовой классификации кожных новообразований с помощью сверточных нейронных сетей. Особое внимание уделяется улучшению показателей метрик классификации для редких кожных поражений, таких, как актинический кератоз (АК), базалиома (ВКЛ), дерматофиброма (DF), сосудистые поражения (VASC) и сквамозно-клеточная карцинома (SCC).

1.1. Данные. В данном исследовании была поставлена задача многоклассовой классификации для дальнейшего анализа и обработки данных. Для проведения поставленных экспериментов были использованы наборы данных, представленные в таблице 1.

Таблица 1. Используемые наборы данных

Название датасета	Количество изображений	Задача	Приватный набор	Метаданные
ISIC-2017 [1], ISIC-2018 [2-3]	15414	Сегментация	-	-
ISIC-2019 [1-2, 4]	33579	Классификация	-	+
ISIC-2020 [5]	42373	Классификация	-	+
MED-NODE [6]	170	Классификация	-	-
SD-198 [7]	6582	Классификация	+	-
7-Point Criteria Database [8]	2019	Классификация	+	+
Light Field Image Dataset of Skin Lesions [9]	114	Классификация	+	-
PH ² [10-11]	200	Классификация	-	+
RAS Skin Dataset	43	Классификация	+	+

Кроме того, набор данных ISIC состоит из других датасетов, таких, как MSK [1], содержащий изображения различных размеров, HAM10000 [2] с изображениями размером 600x450 пикселей и BCN20000 [4] с изображениями размером 1024x1024 пикселей.

Для обучения state-of-the-art сегментационных сверточных нейронных сетей в данном исследовании использовались два датасета: ISIC-2017 [1] и ISIC-2018 [2-3]. Другие датасеты были использованы для обучения классификаторов. В процессе предобработки некоторых изображений выполнялось удаление черных областей, маркировок и других медицинских атрибутов. Кроме того, для предотвращения повторений при обучении использовалось перцептивное хеширование для удаления дубликатов изображений.

Полученный набор данных содержит метаданные пациентов и был сформирован по 9 классам, основываясь на классификации, представленной в датасете ISIC-2019 [1-2, 4]. Любые другие классы кожных поражений, лежащих вне набора классов из ISIC-2019, распределены по иерархической структуре классов заболевания.

Для приватного тестирования моделей была создана отдельная подвыборка из 5066 изображений. Подмножества выборки для обучения, валидации и тестирования были строго определены по индексам в генераторе данных с целью строгого контроля состава каждого подмножества.

Несмотря на относительно большой размер и разнообразие выборки, в ней проявляется значительный дисбаланс классов. Распределение классов в выборке представлено на диаграмме, показанной на рисунке 1.

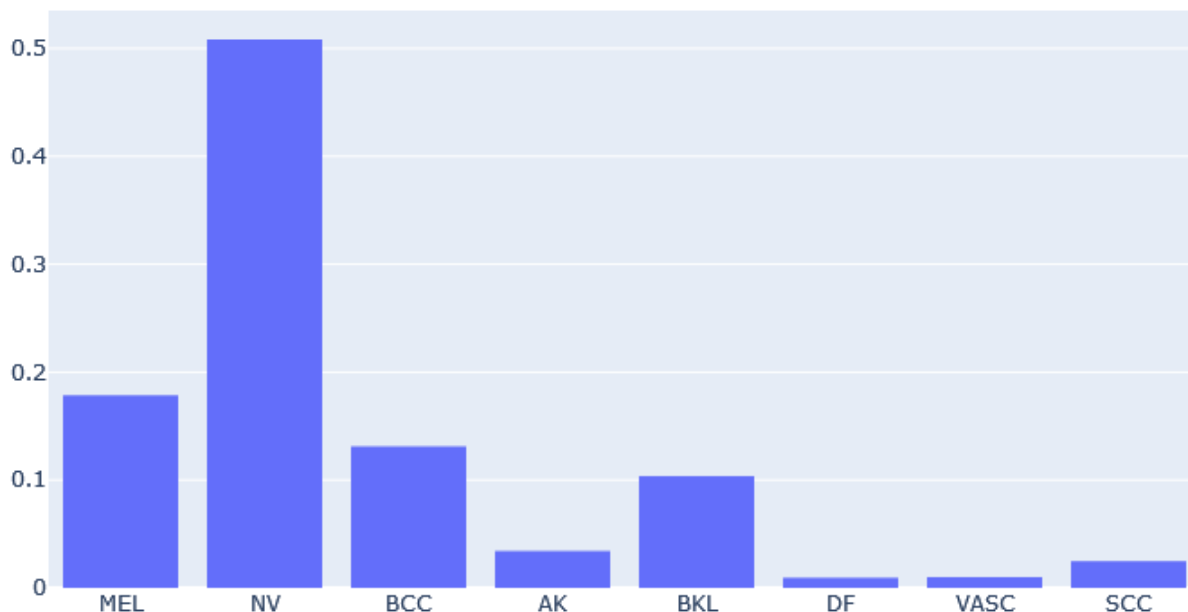


Рис. 1. Распределение классов кожных новообразований

На диаграмме представлены сокращенные названия классов новообразований. Ниже представлен список соответствия классов и их значимые черты:

- меланома (MEL) - участки темного цвета с частичной кровоточивостью;
- меланоцитарный невус (NV) - небольшой выпуклый участок коричневого цвета;
- базальноклеточная карцинома (BCC) - черные базальные клетки;
- актинический кератоз (AK) - участок с выраженным пигментом, покрытый чешуйками;
- кератоз, базалиома (BKL) - темное пигментированное пятно с выраженными краями;
- дерматофиброма (DF) - пятно различного спектра окраски, но преимущественно коричневого цвета;
- сосудистое поражение (VASC) - разрастание кровеносных сосудов;
- сквамозно-клеточная карцинома, спиналиома (SCC) - представляет собой язву, покрытую чешуйчатым налетом.

В списке выше приведено соответствие классов и их значимые особенности, представляющие классификацию из датасета ISIC. Все классы, не вошедшие в этот список и взятые из других датасетов, были распределены в соответствии с иерархической структурой кожных новообразований.

В процессе обучения сверточных нейронных сетей для обеспечения дополнительной регуляризации использовалась аугментация данных. Для тренировочной и валидационной выборок применялись следующие преобразования с определенными вероятностями:

- $p=1$: выполнялись изменение размера 224×224 и нормализация;
- $p=0.5$: применялась операция cutout (1-5) размером не более 16×16 пикселей;
- $p=0.3$: применялось одно из следующих преобразований: ElasticTransform, GridDistortion и OpticalDistortion;
- $p=0.25$: выполнялись повороты, сдвиги, вращения и масштабирования (на 90° , по горизонтали, по вертикали);
- $p=0.15$: применялось одно из следующих преобразований: motion blur и gaussian blur;
- $p=0.05$: выполнялось изменение яркости и контрастности.

Такие преобразования позволяют варьировать и дополнять исходные данные, а также улучшать обобщающую способность модели и значительно снижать вероятность переобучения, обеспечивая регуляризацию модели. Вероятности преобразований были выбраны экспериментально.

1.2. Предлагаемый подход. В данной работе был проведен анализ кожных новообразований с использованием комбинированной архитектуры, состоящей из сегментационных и классификационных моделей сверточных нейронных сетей. Целью комбинированной модели является извлечение дополнительной информации о форме новообразований для решения проблемы классификации редких новообразований. Для классификации кожных новообразований были предложены несколько вариантов комбинированных моделей:

1. Первая модель включает в себя параллельное объединение сегментационной сверточной нейронной сети, классификационной сверточной нейронной сети и полносвязной нейронной сети для учета метаданных пациента. Результаты сегментационной модели подаются через серию $MBConv$ блоков перед объединением, а полученные векторы признаков подаются в несколько линейных слоев обработки.
2. Вторая модель основана на совместном использовании сегментационной и классификационной сверточных нейронных сетей. Результаты сегментационной модели передаются как вход и выход классификационной модели. Аналогично первой модели, результирующая матрица сегментационной модели проходит через серию $MBConv$ блоков перед объединением с результатами классификационной модели, и полученные векторы признаков подаются в несколько линейных слоев обработки.
3. Третья модель реализует комбинированный подход, включающий некоторые свойства предыдущих моделей. Кроме того, в данном варианте добавляется большее количество линейных слоев для каждого выхода модели.

Такой вектор исследований связан с анализом данных, полученных с использованием метода мультиспектральной обработки изображений, при котором данные представляют собой последовательности изображений. Каждый элемент в этих последовательностях формируется путем освещения группами диодов с различными длинами волн. Такие последовательности были получены с использованием специальной камеры, разработанной под руководством Зайченко К. В. ИАП РАН. Благодаря данному методу можно рассматривать кожные новообразования в многомерном пространстве.

В представляемом исследовании данные, полученные с использованием метода мультиспектральной обработки, пока не были использованы, но основной упор сделан на подготовку базы для будущего включения таких данных. Форму новообразования предполагается рассматривать в качестве ключевого фактора.

Первая версия, используемая в исследовании, является базовой моделью, предназначенной для сравнения с последующими улучшениями. В последующих версиях модели была добавлена сеть для извлечения признаков при сегментации.

Во второй версии модели результат работы сверточной нейронной сети для сегментации, обработанный с использованием порога, добавлялся в качестве дополнительного 4-го канала к исходному изображению и подавался на вход классификатора.

Третья версия модели основана на параллельной концепции, в которой исходное изображение обрабатывается параллельно с использованием отдельных сегментационной и классификационной сверточных нейронных сетей.

Четвертая версия модели частично объединяет идеи предыдущих двух версий, используя комбинированный подход для обработки исходного изображения.

2. Обучение и оптимизация моделей. В ходе экспериментов и обучения моделей использовались дополнительные техники для предотвращения переобучения. В первую очередь

были применены L2-регуляризация для обеспечения устойчивости моделей и взвешивание классов с целью учета дисбаланса данных. Оптимизация параметров моделей проводилась путем оценки на кросс-валидации и на валидационной выборке для получения надежных результатов.

Собранный набор данных заранее разделен на тренировочную, валидационную и тестовую выборки. Каждому изображению был присвоен индекс, указывающий его принадлежность к определенной выборке. Размер приватной тестовой выборки составил 5066 образцов. Для каждого образца выполнялись операции нормализации и изменения размера до 224x224.

Для обучения всех моделей в данном исследовании применялись оптимизатор AdamW (с регуляризационным коэффициентом $\text{weight_decay} = 1e-4$) и функция потерь FocalLoss (с $\gamma=2$ и коэффициентом α получена на основе распределения классов). В дополнение к этому проводились эксперименты с использованием оптимизаторов SGD Нестерова с моментумом, равным 0.9 и Adam, в том числе с оптимизацией по кросс-энтропии и кастомными функциями потерь, основанных на кросс-энтропии.

В процессе исследования использовался сервис wandb для сохранения показателей метрик, производительности системы и артефактов модели. Модели оценивались и сохранялись каждые 5 эпох в локальном хранилище, а также один раз в процессе всего обучения в облачном хранилище. Обучение моделей осуществлялось с использованием следующих графических ускорителей: NVIDIA GeForce RTX 4090 (1 шт.), NVIDIA GeForce RTX 3090 Ti (2 шт.), NVIDIA GeForce RTX 3060 (1 шт.) и NVIDIA GeForce RTX 2080 (1 шт.).

2.1. Архитектура SkinNet-v1. Как было отмечено ранее, в качестве базовой модели была использована модель, состоящая из сверточной нейронной сети и полносвязной сети для анализа кожных поражений с учетом метаданных. Последняя представлена на рисунке 2.

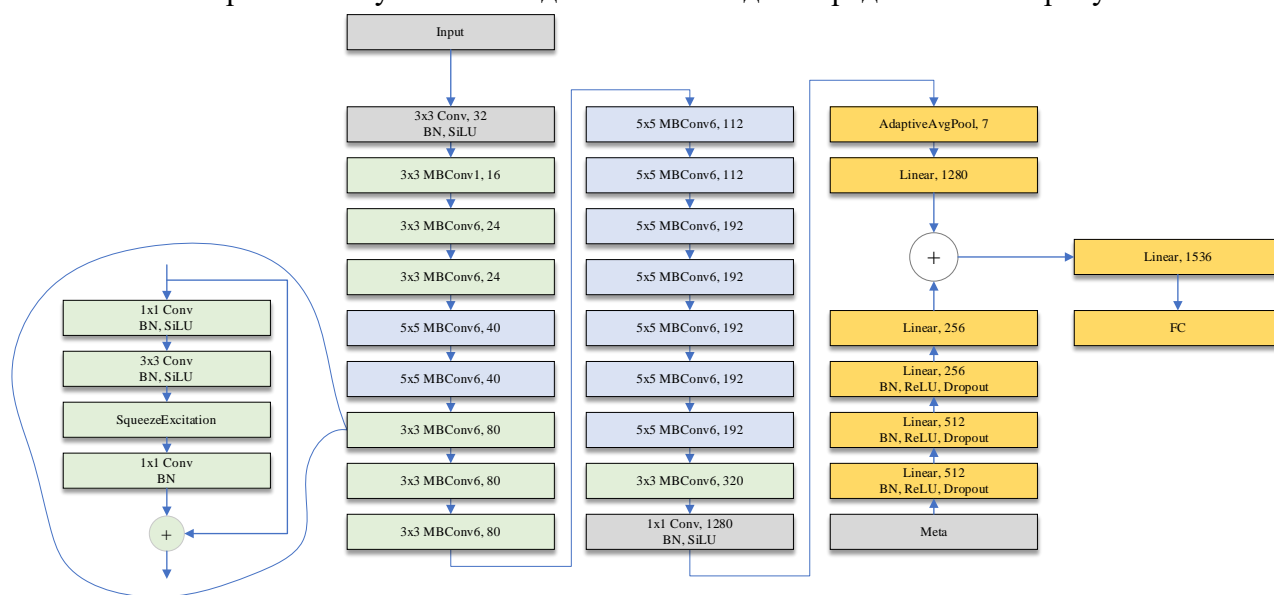


Рис. 2. Архитектура модели SkinNet-v1

В качестве архитектуры сверточной нейронной сети была выбрана модифицированная EfficientNet-B0, продемонстрировавшая высокую эффективность в предыдущих экспериментах и исследованиях, выполненных на аналогичных наборах данных [12]. Модель характеризуется устойчивыми и высокими результатами по ключевым метрикам на этапах кросс-валидации и тестовом наборе данных, а также низкими вычислительными затратами, что стало важным критерием при выборе, учитывая ограниченные вычислительные ресурсы. Помимо EfficientNet-B0, были также оценены такие архитектуры, как ResNet, ResNeXt-101, Inception-ResNet-v2, Inception-ResNet-v3, FixResNext-101, NoisyStudent, MobileNet, SENet154, ViT и ряд других.

После ее обучения веса сверточной нейронной сети были заморожены, затем производилось обучение полносвязной нейронной сети для получения признаков из метаданных. Каждый слой этой сети включает пакетную нормализацию, функцию активации ReLU, в первых трех слоях также применяется Dropout с вероятностью 0.2. Вектор признаков, полученный из нейронной сети по метаданным, объединялся с признаками, полученными из сверточной нейронной сети, и подавался в однослойную сеть с 1536 нейронами, результат которой используется для классификации.

2.2. Архитектура SkinNet-v2. Для решения задачи сегментации был исследован ряд архитектур. На основе метрик, полученных на тестовом наборе данных, и вычислительной сложности была выбрана архитектура Split-Attention U-Net, построенной по архитектурному принципу encoder-decoder. Данная сеть обладает высокими показателями метрик и отличается относительно низкой вычислительной сложностью за счет использования блока Split-Attention в первом слое. В блоке Split-Attention применяются операции свертки Depth-wise Convolution, Point-wise Convolution.

В дополнение к данной архитектуре были проведены эксперименты с несколькими моделями, включая U-Net, R2U-Net, Attention U-Net, Attention R2U-Net, TransUNet, TransCepTion, TransFuse, TransDeepLab, MALUNet, DCSAU-NET и MISSFormer. На тестовом наборе данных наилучшие результаты по метрикам продемонстрировали MALUNet (mIoU = 0.79759) и MISSFormer (mIoU = 0.78715). Модель Split-Attention U-Net (mIoU = 0.77408) продемонстрировала сопоставимые показатели метрик, однако значительно превосходит другие архитектуры по скорости обучения благодаря последовательному применению свёрточных блоков Depth-wise и Point-wise. Операция Depth-wise свертки выполняется независимо для каждого входного канала, используя ядра размером 5x5, что уменьшает вычислительную сложность, так как применяет свертку только по глубине входного тензора, отдельно для каждого канала. Операция Point-wise свертки с использованием ядра размером 1x1 выполняется по всем пространственным размерам входного тензора после Depth-wise свертки для агрегации признаков с разных каналов входного тензора.

Каждый последующий encoder-блок завершается операцией max pooling размером 2x2, в то время как каждый decoder-блок начинается с операции upsampling, которая последовательно восстанавливает оригинальный размер входного фрейма. Skip connections применяются для объединения признаков (feature maps) из блока декодера с соответствующими признаками из блока энкодера, что способствует интеграции низкоуровневой и высокоуровневой информации. Завершающий слой модели представляет собой свертку 1x1, что позволяет формировать бинарные или многоклассовые маски для сегментации кожных новообразований.

На рисунке 3 сегментационная сеть обозначена красной областью. Входное изображение подается на вход сегментационной сети. Затем результат конкатенируется с входным изображением в виде четвертого канала и подается на вход классификатора SkinNet-v1. При объединении к результату сегментационной сверточной нейронной сети применяется пороговая функция, которая формирует бинарную маску для кожного поражения.

По показателям метрик данная модель работает хуже, однако количество неправильно предсказанных редких поражений уменьшилось на 4%.

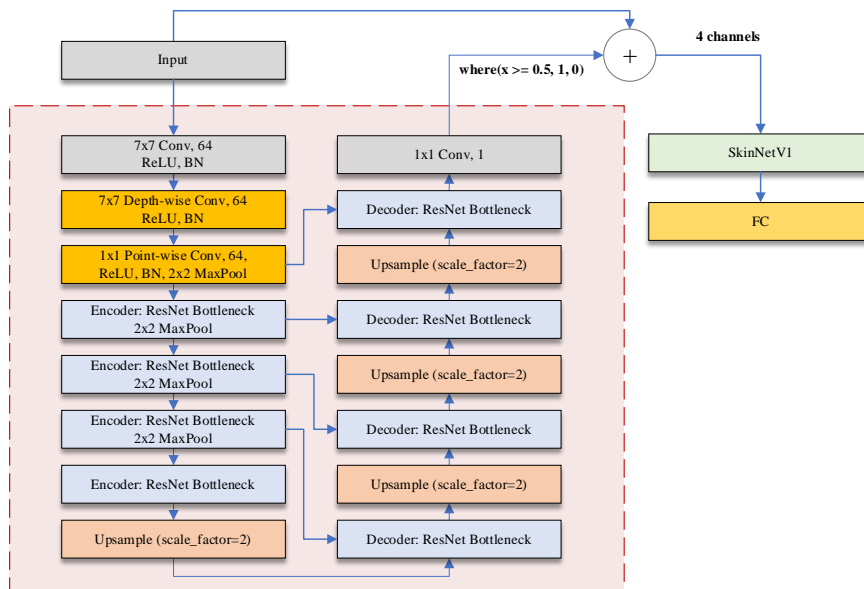


Рис. 3. Архитектура модели SkinNet-v2

2.3. Архитектура SkinNet-v3. На рисунке 4 представлена третья версия модели, отличительной особенностью которой является внедрение небольшой сверточной сети для обработки вывода сегментационной сети, обозначенной фиолетовой областью.

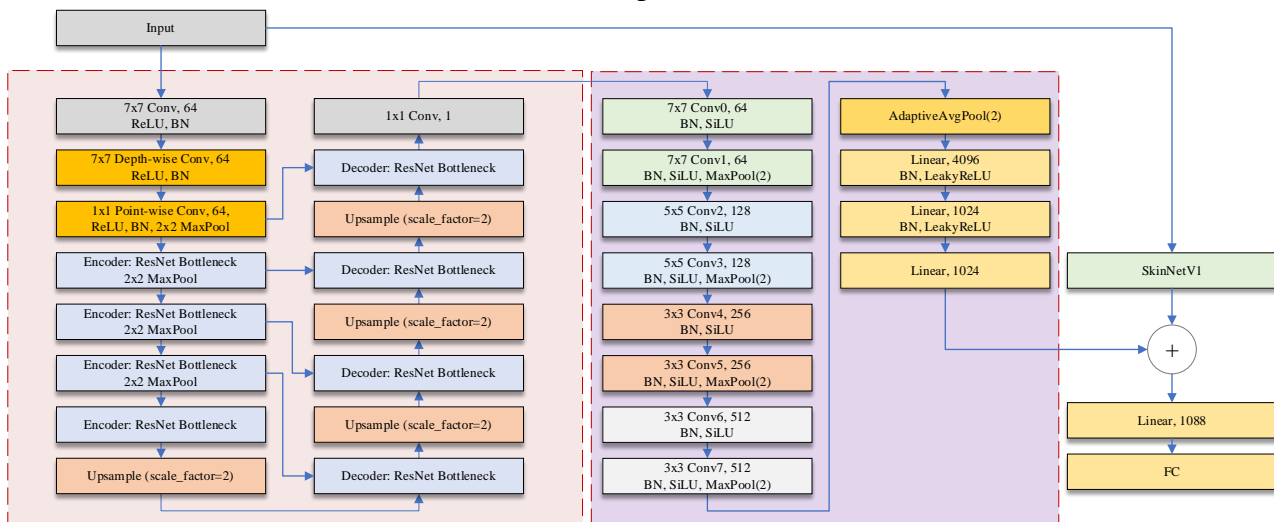


Рис. 4. Архитектура модели SkinNet-v3

Исходное входное изображение подается на вход классификационной сети, а результат конкатенируется с 64 признаками из блока сегментационной сети. Метрики данной модели незначительно, но продемонстрировали улучшение по сравнению с предыдущими моделями. Результат сегментационной сети, обозначенной светло-красным областью, в данном случае представляет из себя изображение поражения, которое было выделено с использованием бинарной маски, аналогично семантической сегментации.

Классификационная сверточная сеть, предназначенная для обработки вывода сегментационной сети, состоит из двух сверток 7x7, двух сверток 5x5, четырех сверток 3x3 и нескольких линейных слоев. После каждого сверточного слоя применяется операция MaxPool 2x2. Результат сети, обозначенной фиолетовой областью, является вектором признаков размерности 1024. Исходное входное изображение подается на вход классификационной сети, а результат конкатенируется с 64 признаками из блока сегментационной сети. Показатели метрик данной модели незначительно, но продемонстрировали улучшение по сравнению с предыдущими моделями.

2.4. Архитектура SkinNet-v4. Данная модель является наиболее сложной в плане вычислительной мощности в серии проведенных экспериментов. С точки зрения концепции, она является комбинацией второй и третьей версий моделей (рис.5).

Однако вместо дублирования модели первой версии для обработки результатов сегментации была добавлена более модифицированная сверточная сеть на основе EfficientNet-B3. Было замечено, что метрики улучшились в среднем на 1% по сравнению с базовой версией.

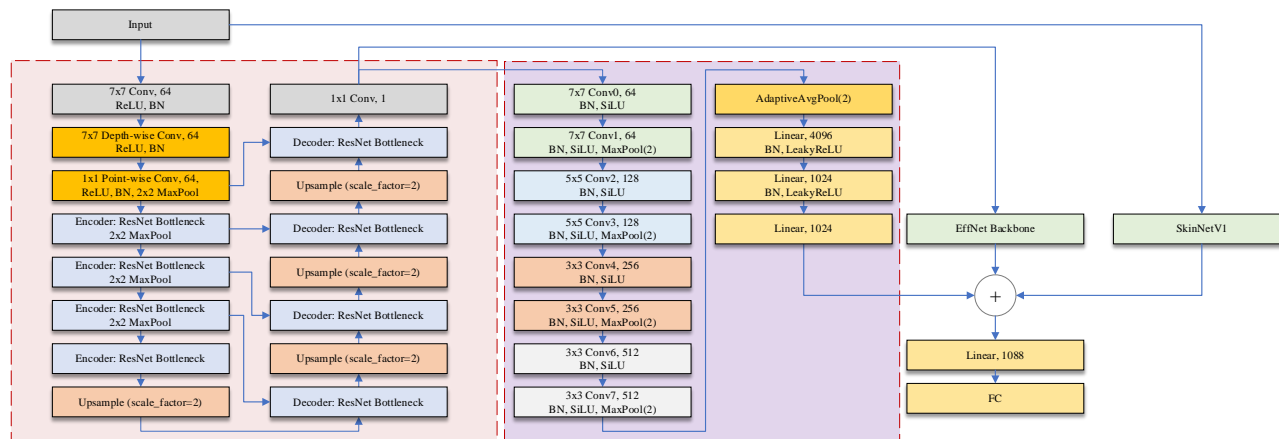


Рис. 5. Архитектура модели SkinNet-v4

3. Оценка моделей. Стоит отметить, что в данном исследовании, помимо тренировочной и валидационной выборок, была сформирована тестовая выборка, состоящая из 5066 изображений, которая моделирует оценку процесса эксплуатации моделей на новых данных. Тестовая выборка аналогичным образом строго предопределялась по индексам. В задаче классификации оценка моделей выполнялась с использованием матрицы ошибок и взвешенных метрик Recall, Precision и F1-score, а также обычных метрик Specificity, Recall, Precision и F1-score. Для лучшего понимания внесенных изменений в архитектуру моделей в результатах приведены расширенные расчеты метрик для каждого из классов. Для оценки сегментационных моделей использовались метрики IoU и Dice, результаты которых лежат вне рамок данного исследования.

Модели, в том числе и ансамбли, были оценены на тестовом наборе данных, а также с использованием стратифицированной кросс-валидации. Кроме того, в данном исследовании особое внимание уделялось оценке суммарных ошибок первого и второго рода на основе матрицы ошибок классификации как для всех типов новообразований, так и для редких кожных поражений (AK, BKL, DF, VASC, SCC).

Результаты модели SkinNet-v1 представлены в таблице 2. Суммарное значение FP+FN для всех новообразований составляет 655, в то время как для редких поражений это значение равно 199. Эти значения будут использоваться в качестве отправной точки для дальнейшего анализа. Результаты модели SkinNet-v2 представлены в таблице 3. Суммарное значение FP+FN для всех типов новообразований составляет 676, что на 3% меньше по сравнению с моделью SkinNet-v1. Для редких поражений значение FP+FN составляет 191. Несмотря на то, что устойчивая к дисбалансу классов метрика Recall улучшилась, взвешенные метрики показывают обратную тенденцию.

Таблица 2. Показатели метрик модели SkinNet-v1

Class	Specificity	Recall	Precision	F1-Score	WRecall	WPrecision	WF1
MEL	0.9586	0.8374	0.8067	0.8217	0.1433	0.1381	0.1406
NV	0.9359	0.9328	0.9410	0.9368	0.4876	0.4918	0.4897
BCC	0.9889	0.8294	0.9263	0.8752	0.1190	0.1329	0.1256
AK	0.9851	0.5789	0.5756	0.5773	0.0195	0.0194	0.0195

BKL	0.9673	0.8004	0.7028	0.7484	0.0705	0.0619	0.0659
DF	0.9982	0.7826	0.7826	0.7826	0.0071	0.0071	0.0071
VASC	0.9988	0.8491	0.8824	0.8654	0.0089	0.0092	0.0091
SCC	0.9923	0.6944	0.6637	0.6787	0.0148	0.0141	0.0145
Total	0.9781	0.7881	0.7851	0.7858	0.8707	0.8746	0.8719

Таблица 3. Показатели метрик модели SkinNet-v2

Class	Specificity	Recall	Precision	F1-Score	WRecall	WPrecision	WF1
MEL	0.9554	0.8186	0.7922	0.8052	0.1407	0.1362	0.1384
NV	0.9293	0.9347	0.9354	0.9350	0.4886	0.4889	0.4887
BCC	0.9871	0.8248	0.9137	0.8670	0.1171	0.1297	0.1230
AK	0.9849	0.6045	0.5912	0.5978	0.0211	0.0207	0.0209
BKL	0.9710	0.7652	0.7167	0.7402	0.0669	0.0627	0.0647
DF	0.9976	0.8261	0.7451	0.7835	0.0075	0.0068	0.0071
VASC	0.9978	0.8679	0.8070	0.8364	0.0091	0.0084	0.0087
SCC	0.9939	0.7248	0.7248	0.7248	0.0156	0.0156	0.0156
Total	0.9771	0.7958	0.7783	0.7862	0.8666	0.8689	0.8673

Результаты модели SkinNet-v3 представлены в таблице 4. Суммарное значение FP+FN для всех типов новообразований составляет 625 (улучшение на 7,5% по сравнению с предыдущей моделью), а для редких поражений – 164.

Таблица 4. Показатели метрик модели SkinNet-v3

Class	Specificity	Recall	Precision	F1-Score	WRecall	WPrecision	WF1
MEL	0.9562	0.8053	0.7916	0.7984	0.1380	0.1356	0.1368
NV	0.9321	0.9490	0.9388	0.9439	0.4962	0.4909	0.4935
BCC	0.9894	0.8456	0.9297	0.8857	0.1200	0.1319	0.1257
AK	0.9877	0.5954	0.6319	0.6131	0.0203	0.0216	0.0209
BKL	0.9762	0.8000	0.7660	0.7826	0.0711	0.0680	0.0695
DF	0.9968	0.8043	0.6981	0.7475	0.0073	0.0063	0.0068
VASC	0.9980	0.7547	0.8000	0.7767	0.0079	0.0084	0.0081
SCC	0.9929	0.7407	0.6957	0.7175	0.0158	0.0148	0.0153
Total	0.9787	0.7869	0.7815	0.7832	0.8766	0.8776	0.8767

При сравнении первых трех моделей по метрике Recall можно предположить, что вторая модель является наиболее точной, в то время как третья модель имеет наименьшее значение данной метрики. Однако при расчете взвешенных метрик становится ясно, что лучшей моделью является SkinNet-v3. Если сравнивать модели по сумме ошибок первого и второго рода, можно сделать вывод, что третья модель имеет наименьшее количество ошибок, как для всех типов новообразований, так и для редких поражений.

Результаты модели SkinNet-v4 представлены в таблице 5. Суммарное значение FP+FN для всех типов новообразований составляет 561, что является улучшением на 10% по сравнению с предыдущей моделью, в то время как для редких поражений значение FP+FN составляет 130.

Таблица 5. Показатели метрик модели SkinNet-v4

Class	Specificity	Recall	Precision	F1-Score	WRecall	WPrecision	WF1
MEL	0.9612	0.8088	0.8116	0.8102	0.1386	0.1391	0.1388
NV	0.9368	0.9539	0.9429	0.9484	0.4986	0.4928	0.4957
BCC	0.9894	0.8695	0.9323	0.8998	0.1250	0.1340	0.1293
AK	0.9886	0.6491	0.6647	0.6568	0.0219	0.0224	0.0222
BKL	0.9766	0.8184	0.7717	0.7943	0.0720	0.0679	0.0699

DF	0.9988	0.8696	0.8696	0.8696	0.0079	0.0079	0.0079
VASC	0.9992	0.8491	0.9000	0.8738	0.0089	0.0094	0.0091
SCC	0.9952	0.7870	0.7798	0.7834	0.0168	0.0166	0.0167
Total	0.9807	0.8257	0.8341	0.8295	0.8897	0.8902	0.8897

В ходе экспериментов подобных результатов удалось достичь путем параллельного использования модели EffcientNet-V3 в качестве второго классификатора и с модификацией некоторых сверточных слоев с их последующим дообучением.

4. Результаты. В таблице 6 представлены показатели метрик, полученные разработанными моделями на тестовом наборе данных. Напомним, что основной целью нашего исследования является улучшение показателей метрик при классификации редких новообразований.

Таблица 6. Общие показатели метрик

Model	All: FP+FN	Rares: FP+FN	Specificity	WRe-call	WPrecision	WF1-score
SkinNet-v1	655	199	0.9781	0.8707	0.8746	0.8719
SkinNet-v2	676	191	0.9771	0.8666	0.8689	0.8673
SkinNet-v3	625	164	0.9787	0.8766	0.8776	0.8767
SkinNet-v4	561	130	0.9807	0.8897	0.8902	0.8897

Согласно результатам, наиболее эффективной для достижения этой цели оказалась четвертая версия модели. Данная модель не только снизила количество неверных классификаций редких новообразований, но также привела к значительному снижению числа ошибок для всех классов и, как следствие, к улучшению всех представленных метрик.

Из таблицы видно, что каждая последующая версия модели демонстрировала лучший потенциал в анализе редких новообразований. Например, SkinNet-v4 показывает лучшие показатели метрик с точки зрения Specificity и взвешенных Recall, Precision и F1-score. Кроме того, данная модель характеризуется наименьшим количеством ложноположительных и ложноотрицательных результатов как для всех классов, так и для редких образцов.

Заключение. В данном исследовании был представлен метод использования комбинированной модели с расширенным извлечением дополнительной информации о форме кожных новообразований с целью решения проблемы анализа редких новообразований. Тем самым, это позволяет более точно классифицировать эти редкие кожные поражения.

Был проведен ряд экспериментов, направленных на минимизацию ошибок первого и второго рода для редких кожных заболеваний (AK, BKL, DF, VASC, SCC).

На основе этих экспериментов были сформулированы четыре версии модели, приведенные в данной статье.

В дальнейших исследованиях планируется:

- проверить гипотезу исследования в других медицинских областях;
- оценить влияние генерации изображений редких классов кожных поражений с помощью модели на основе генеративно-состязательных сетей и операции MixUp на показатели метрик классификации базовой модели;
- проверить гипотезу исследования на лабораторных крысах с искусственно выведенными кожными новообразованиями.

Все это позволит улучшить показатели метрик моделей для классификации медицинских заболеваний, обученных на данных с сильным дисбалансом данных.

Благодарности. Работа была поддержана Минобрнауки РФ, госзадание № 075-01157-23-00, тема № FZZM-2022-0011.

Список источников

1. Codella N. et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI). IEEE 15th International symposium on biomedical imaging (ISBI 2018), 2018, pp. 168-172, DOI: 10.1109/ISBI.2018.8363547
2. Tschandl P., Rosendahl C., Kittler H. The HAM10000 dataset, a large collection of multi-sources dermatoscopic images of common pigmented skin lesions. Sci Data 5, 180161, 2018, DOI: 10.1038/sdata.2018.161.
3. Codella N., Rotemberg V., Tschandl P. et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC), 2018, arXiv: 1902.03368 [cs.CV]
4. Combalia M., Codella N., Rotemberg V., et al. BCN20000: Dermoscopic lesions in the wild, arXiv:1908.02288, 2019, DOI: 10.48550/arXiv.1908.02288
5. Rotemberg V., Kurtansky N., Betz-Stablein B., et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci Data 8, 34. 2021, DOI: 10.1038/s41597-021-00815-z
6. Giotis I., Molders N., Land S., Biehl M., Jonkman M.F., Petkov N. MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert systems with applications, 2015, vol. 42, pp. 6578-6585.
7. Sun X., Yang J., Sun M., Wang K. A benchmark for automatic visual classification of clinical skin disease images. ECCV: Springer, 2016, vol. 9910, pp. 206–222.
8. Kawahara J., Daneshvar S., Argenziano G., Hamarneh G. Seven-point checklist and skin lesion classification using multi-task multi-modal neural nets. IEEE Journal of biomedical and health informatics (IEEE JBHI) special issue on Skin lesion image analysis for Melanoma Detectio, 2019, vol. 23, no. 2, pp. 538-546, DOI: 10.1109/JBHI.2018.2824327.
9. Faria S.M.M., et al. Light field image dataset of skin lesions. IEEE 41st Annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2019, pp. 3905–3908, DOI:10.1109/EMBC.2019.8856578.
10. Barata C., Ruela M., et al. Two systems for the detection of melanomas in Dermoscopy images using texture and color features. IEEE systems journal, 2013, 8(3), pp. 965–979.
11. Mendonça T., Ferreira P. M., Marques J., Marcal A., Rozeira J. PH² - A dermoscopic image database for research and benchmarking. IEEE 35th international conference of the IEEE engineering in medicine and biology society, Osaka, Japan, 2013, DOI:10.1109/EMBC.2013.6610779.
12. Milantev S., Bessmertnyi I., et al. Skin lesion analysis using ensemble of CNN with dermoscopic images and metadata. CEUR workshop proceedings, 2021, vol. 2893.

Милантьев Сергей Андреевич. Аспирант факультета Программной инженерии и компьютерной техники университета ИТМО, младший научный сотрудник ИАП РАН. AuthorID: 1176348, SPIN: 5509-8366, ORCID: 0000-0002-1970-5217, nevdev@niuitmo.ru, Россия, Санкт-Петербург, Кронверкский проспект, 49

Святкина Виталия Игоревна. Младший научный сотрудник ИАП РАН, ORCID: 0000-0002-6456-8776, svyatkina.vi@edu.spbstu.ru, Россия, Санкт-Петербург, Ивана Черных, 31-33

Бессмертный Игорь Александрович. Профессор, доктор технических наук, университет ИТМО. AuthorID: 571032, SPIN: 4329-6492, ORCID: 0000-0001-6711-6399, bessmertny@itmo.ru, Россия, Санкт-Петербург, Кронверкский проспект, 49

Зайченко Кирилл Вадимович. Профессор, доктор технических наук, зав. лаб., главный научный сотрудник ИАП РАН. AuthorID: 107276, SPIN: 7951-3906, ORCID: 0000-0002-2881-4386, Россия, Санкт-Петербург, Ивана Черных, 31-33

UDC: 004.93.12

DOI:10.25729/ESI.2024.35.3.004

Image analysis of skin lesion using a combined convolutional neural network architecture

Sergey A. Milantev^{1,2}, Vitaliia I. Sviatkina², Igor A. Bessmertny¹, Kirill V. Zaichenko²

¹ITMO University, Russia, Saint Petersburg, nevdev@niuitmo.ru

²IAI RAS, Russia, Saint Petersburg

Abstract. This research explores the possibility of applying combined convolutional neural network architectures to analyze skin lesions. Model architectures have been developed to extract additional features related to the shape pattern of skin lesions. Optimization of the models, including the architecture, was performed in order to minimize

I and II types of errors for rare skin lesions. ISIC2017-2020, MED-NODE, SD-198, 7-Point Criteria Database, Light Field Image Dataset of Skin Lesions, PH2, IAP RAS were used in the training process. AdamW optimizer, FocalLoss functions and CosineAnnealingWarmRestarts scheduler were used to train classification models. The BCEDice loss function was used to train the segmentation models. The models were evaluated using weighted classification metrics such as Specificity, Recall, Precision and F1-score. The robustness of model architecture was considered during the validation phase. Models which are using additional convolutional neural networks for the skin lesion extraction shape features showed better metrics performance and also had lower sum of I and II type errors for rare lesions compared to conventional classification models. The results of this research can be used in analyzing medical problems with data imbalance in the training dataset.

Keywords: skin lesions, convolutional neural networks, skin lesion analysis, class imbalance, multispectral image processing

Acknowledgements: This work was supported by the Ministry of Science and Higher Education of the Russian Federation, state task no. 075-01157-23-00, topic no. FZZM-2022-0011.

References

1. Codella N. et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI). IEEE 15th International symposium on biomedical imaging (ISBI 2018), 2018, pp. 168-172, DOI: 10.1109/ISBI.2018.8363547
2. Tschandl P., Rosendahl C., Kittler H. The HAM10000 dataset, a large collection of multi-sources dermatoscopic images of common pigmented skin lesions. Sci Data 5, 180161, 2018, DOI: 10.1038/sdata.2018.161.
3. Codella N., Rotemberg V., Tschandl P. et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC), 2018, arXiv: 1902.03368 [cs.CV]
4. Combalia M., Codella N., Rotemberg V., et al. BCN20000: Dermoscopic lesions in the wild, arXiv:1908.02288, 2019, DOI: 10.48550/arXiv.1908.02288
5. Rotemberg V., Kurtansky N., Betz-Stablein B., et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci Data 8, 34. 2021, DOI: 10.1038/s41597-021-00815-z
6. Giotis I., Molders N., Land S., Biehl M., Jonkman M.F., Petkov N. MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert systems with applications, 2015, vol. 42, pp. 6578-6585.
7. Sun X., Yang J., Sun M., Wang K. A benchmark for automatic visual classification of clinical skin disease images. ECCV: Springer, 2016, vol. 9910, pp. 206–222.
8. Kawahara J., Daneshvar S., Argenziano G., Hamarneh G. Seven-point checklist and skin lesion classification using multi-task multi-modal neural nets. IEEE Journal of biomedical and health informatics (IEEE JBHI) special issue on Skin lesion image analysis for Melanoma Detectio, 2019, vol. 23, no. 2, pp. 538-546, DOI: 10.1109/JBHI.2018.2824327.
9. Faria S.M.M., et al. Light field image dataset of skin lesions. IEEE 41st Annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2019, pp. 3905–3908, DOI:10.1109/EMBC.2019.8856578.
10. Barata C., Ruela M., et al. Two systems for the detection of melanomas in Dermoscopy images using texture and color features. IEEE systems journal, 2013, 8(3), pp. 965–979.
11. Mendonça T., Ferreira P. M., Marques J., Marcal A., Rozeira J. PH² - A dermoscopic image database for research and benchmarking. IEEE 35th international conference of the IEEE engineering in medicine and biology society, Osaka, Japan, 2013, DOI:10.1109/EMBC.2013.6610779.
12. Milantev S., Bessmertnyi I., et al. Skin lesion analysis using ensemble of CNN with dermoscopic images and metadata. CEUR workshop proceedings, 2021, vol. 2893.

Milantev Sergey Andreevich. PhD student of the faculty of software engineering and computer systems, ITMO university, junior researcher at IAI RAS, AuthorID: 1176348, SPIN: 5509-8366, ORCID: 0000-0002-1970-5217, nevdev@niuitmo.ru, Russia, St. Petersburg, Kronverksky Prospect, 49

Sviatkina Vitaliia Igorevna. Junior researcher at IAI RAS, ORCID: 0000-0002-6456-8776, sviatkina.vi@edu.spbstu.ru, Russia, St. Petersburg, Ivan Chernykh, 31-33

Bessmertny Igor Alexandrovich. Professor, Doctor of technical science, AuthorID: 571032, SPIN: 4329-6492, ORCID: 0000-0001-6711-6399, bessmertny@itmo.ru, Russia, St. Petersburg, Kronverksky Prospect, 49

Zaichenko Kirill Vadimovich. Professor, doctor of technical sciences, head of the laboratory at IAI RAS, AuthorID: 107276, SPIN: 7951-3906, ORCID: 0000-0002-2881-4386, Russia, St. Petersburg, Ivan Chernykh, 31-33

Статья поступила в редакцию 18.10.2023; одобрена после рецензирования 11.10.2024; принята к публикации 11.10.2024.

The article was submitted 10/18/2023; approved after reviewing 10/11/2024; accepted for publication 10/11/2024.

УДК 004.891

DOI:10.25729/ESI.2024.35.3.005

Экспертная система поддержки расследования происшествий на железнодорожном транспорте

Подобедова Елизавета Вячеславовна, Жукова Марина Сергеевна,
Аршинский Леонид Вадимович

Иркутский государственный университет путей сообщения,
Россия, Иркутск, *larsh@mail.ru*

Аннотация. В работе рассмотрен подход к разработке экспертной системы по анализу происшествий на железнодорожном транспорте. Обсуждаются этапы и особенности разработки базы знаний, возможности инструментальной среды разработки и другие вопросы. Прототип системы разработан для анализа сходов, однако принятая в ней (продукционная) модель знаний позволяет относительно легко расширить систему и на другие обстоятельства происшествий. Система реализована в среде, использующей вывод на основе логик с векторной семантикой, которые позволяют работать как с достоверной, так и с нечёткой, неопределённой и противоречивой информацией.

Ключевые слова: экспертная система, векторная семантика, расследование происшествий, железнодорожный транспорт

Цитирование: Подобедова Е.В. Экспертная система поддержки расследования происшествий на железнодорожном транспорте / Е.В. Подобедова, М.С. Жукова, Л.В. Аршинский // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 57-72. – DOI: 10.25729/ESI.2024.35.3.005.

Введение. Российские железные дороги – один из важнейших компонентов транспортного комплекса Российской Федерации. Железная дорога оказывает огромное влияние на экономическое благополучие государства, оказывает воздействие на формирование его федерального и регионального бюджетов и также играет важное значение в транспортной логистике. Железная дорога считается главным межрегиональным видом транспорта в пределах экономически освоенной территории страны.

Вопросы эффективности функционирования транспортного комплекса и развития транспортной инфраструктуры играют важную роль в развитии российской экономики по причине обширной территории страны и больших расстояний между центрами производства и центрами потребления [1]. Повышение эффективности может, с одной стороны, привлечь дополнительные транзитные грузопотоки, а с другой – повысить конкурентоспособность российской экономики и потенциал для привлечения дополнительных инвестиций. Одним из препятствий к росту эффективности перевозок являются происшествия на железнодорожном транспорте, которые замедляют его работу и ведут к дополнительным потерям при перевозках. Последствия от происшествий можно ослабить, если повысить качество их расследования и сократить время его проведения.

Повышения качества и сокращения сроков расследования можно добиться привлечением компетентных специалистов и автоматизацией отдельных его этапов [2]. Одним из подходов здесь является технология экспертных систем (ЭС). Подобные системы содержат в себе высокопрофессиональные знания и позволяют принимать решения или, как минимум, выдвигать и проверять рабочие версии по соответствующему вопросу [3, 4].

Важным применением ЭС может стать обучение соответствующих специалистов [5, 6].

1. Описание предметной области. Железная дорога в широком смысле – это комплексное предприятие, располагающее всеми техническими средствами, обслуживаемыми специалистами, осуществляющими эксплуатацию этих средств и организующими перевозки пассажиров и грузов [1]. Железная дорога в узком смысле – это дорога с рельсовыми путями, пред-

назначенная для передвижения железнодорожного подвижного состава. Сюда входит оборудованная рельсами полоса земли либо поверхности искусственного сооружения (тоннель, мост, эстакада), которые используются для движения рельсовых транспортных средств.

Железные дороги разделяются на магистральные, промышленные (подъездные пути предприятий и организаций) и городские (метрополитен, трамвай). По числу путей они делятся на одна-, двух- и многопутную. Прибытие подвижного состава на станцию или отправление со станции, контроль его передвижения, составление расписания подвижного состава, регулировка скорости поезда и т.д. требуют большой ответственности целой команды специалистов. Сложность этой системы приводит к тому, что, к сожалению, на железнодорожном транспорте иногда происходят чрезвычайные происшествия. Они обусловлены разными причинами и разделяются (классифицируются) на [7]:

- столкновения поездов или железнодорожного подвижного состава с другими поездами или подвижным составом;
- сходы поездов (на перегоне или железнодорожной станции);
- столкновения с автотракторной техникой;
- происшествия при перевозке опасных грузов;
- развал грузов;
- прием или отправление поезда по неготовому маршруту;
- проезд запрещающего сигнала светофора или предельного столбика;
- приём поезда на занятый путь или отправление на занятый перегон;
- разрушение конструктивных элементов вагона или состава (излом оси, осевой шейки или колеса, излом боковины или надрессорной балки тележки вагона, обрыв хребтовой балки подвижного состава);
- несанкционированное движение по железнодорожным путям транспортного средства или столкновение железнодорожного подвижного состава с транспортным средством вне установленного железнодорожного переезда;
- перевод стрелки под железнодорожным подвижным составом;
- ложное появление на светофоре разрешающего показания сигнала вместо запрещающего или появление более разрешающего показания сигнала вместо показания, требующего продолжения следования поезда с уменьшенной скоростью;
- и т.д.

Одним из наиболее распространённых происшествий является сход [8, 9]. Сходом железнодорожного подвижного состава называется событие, при котором хотя бы одно колесо сошло с головки рельса и для его постановки на рельс требуется применение подъемных средств и (или) приспособлений [10]. Сходы делятся (классифицируются) на следующие виды [8]:

1. Сход из-за распора колеи. Это уширение колеи от её нормы (1520мм +8;-4). Если колея будет больше, а поезд поедет по установленной скорости равной 60-80 км/ч, то колеса подвижной состав могут сойти на середину пути.

2. Сход из-за выжимания экипажа. Экипаж – это нижняя часть вагона, локомотива, тележки и т.д., т.е. всего, что ниже кузова. Часто выжимание экипажа происходит из-за неправильно сформированного состава, когда груженные вагоны ставят в хвост, а порожние в голову.

3. Сход из-за недопустимого горизонтального поперечного искривления (сдвига) колеи колесами «шальной» тележки при торможении поезда или из-за ее температурного выброса. Этот сход характерен для летнего сезона. Есть стыковой и бесстыковой путь и есть температурные рельсовые зазоры. Допустимая норма рельсовых зазоров 20-22 мм. Если в рельсовом

стыке возник слепой зазор, т.е. нарушились допустимые нормы, происходит выброс рельса; соответственно, идет уширение колеи и дальнейший сход подвижного состава.

4. Сход из-за излома рельса. Излом рельса характерен для низких температур, когда зазор превышает 26-28 мм. В этом случае либо разрывает сам стык, либо рельс в любой его части.

5. Сход из-за накатывания на неприлегающее к рамному рельсу остриё остряка стрелочного перевода, подрезанного гребня «шальной» тележки и вползания гребня на рамный рельс и остряк. Если остряк стрелочного перевода, прилегание которого к рамному рельсу не должно превышать 4 мм, неплотно прилегает к рельсу, то может произойти сход. В холодное время года это может случиться при налипании льда, летом – при попадании грязи, которая вовремя не убирается. Происшествие в основном характерно для т.н. «противошерстного» движения, когда колёса едут на остряк.

6. Сходы в кривых из-за избыточного возвышения наружного рельса. Кривая лежит под углом и имеет внутренний и наружный рельс. Если наружный рельс возвышается над внутренним более 75 мм, то может произойти два развития событий: либо колесная пара уйдет в сторону наружного колеса кривой, либо гребень колеса наткнется на головку рельса.

Сход железнодорожного подвижного состава также разделяется на:

1. Сход на перегонах: данный тип схода относят к авариям на железнодорожном транспорте, т.к. заваливается состав на перегоне, перекрываются пути и приходится восстанавливать их, что влечет за собой материальный ущерб. Это наиболее редкие сходы.

2. Сход на станциях: происходит, когда состав выезжает/въезжает на станцию, а также на самой станции, при маневровых работах. В данном случае состав не заваливается, а сходит тележка, которую необходимо затащить. Эти сходы наиболее часто встречаются там, где есть сортировочно-подгоночные парки.

К сходу железнодорожного подвижного состава приводят разные факторы: нарушения технологии загрузки/разгрузки вагона, особенности маршрута, несвоевременный перевод стрелки, сейсмические события и др. Но основные факторы, влияющие на сход, – это состояние пути и технические параметры вагона [9] (иные причины, вроде столкновения поездов, столкновения с автотракторной техникой, теракт, неуправляемое движение и т.д. здесь не рассматриваются, т.к. это отдельные виды происшествий, требующие своего расследования).

2. Экспертная система анализа происшествий на железнодорожном транспорте.

Предназначение предлагаемой ЭС – консультирование по причинам возникновения схода; также она может способствовать повышению квалификации работников железнодорожного транспорта и иных лиц, привлекаемых к расследованию. Главным достоинством ЭС является возможность неограниченного накопления и хранения знаний в системе. В качестве экспертов, при помощи которых будут накапливаться знания в БЗ, выступает ревизорский аппарат, состоящий из ревизоров службы движения, электроснабжения, инфраструктуры, СЦБ и связи, а также ревизора локомотивной бригады. Могут привлекаться и иные лица, имеющие опыт расследований.

Основными функциональными возможностями ЭС являются:

- проверка гипотез о причине возникновения происшествий;
- прогнозирование возможности происшествий на основе имеющихся факторов;
- сохранение опыта специалистов путём занесения его в БЗ.

В связи с тем, что БЗ должна быть расширяемой (это позволит ЭС оставаться открытой для усовершенствования и обновления информации), целесообразно выбрать продукционную модель знаний, главным качеством которой является простота редактирования БЗ.

Функционировать ЭС должна как в режиме консультанта: помогать в поиске причин и факторов возникновения происшествия и формировать рабочие гипотезы, так и в режиме приобретения знания: пополнять свою БЗ за счет новых фактов и правил, полученных от экспертов и из специальной литературы.

3. Проектирование базы знаний. Необходимым компонентом ЭС является БЗ, которая содержит в себе модель предметной области. БЗ должна быть структурирована и специальным образом представлена [11]. Наиболее подходящей формой представления знаний для выбранной предметной области является продукционная, в основе которой лежит конструкция «Если..., то...». При этом левая часть правила продукции (антецедент) здесь может содержать связки И, ИЛИ, НЕ, а также скобки. С точки зрения авторов это облегчает восприятие БЗ и сокращает её объём.

Основным достоинством продукционной модели знаний является простота модификации БЗ, построенной на её основе. Единицы знаний таких БЗ интуитивно понятны, легко дополняются новыми, могут удаляться, переписываться, и все это – не затрагивая других единиц знаний. Важное условие здесь – сохранение непротиворечивости системы знаний. Противоречие может сделать систему неработоспособной в соответствующих ситуациях или ослабит доказательную силу знаний (при правдоподобном выводе). Существуют и другие сложности [12-15]. Однако это не снижает общий интерес к подобным моделям. Вопросы, связанные с поиском и устранением противоречий БЗ в рамках используемой далее семантики достаточно подробно изложены в [16, 17].

Одним из практических способов, с помощью которого в БЗ могут формироваться единицы знаний, взаимосвязи предметной области, является диаграмма Исикавы. Это графический метод отображения и анализа причинно-следственных связей, который позволяет систематизировать причины тех или иных событий в виде специального графического представления, называемого «рыбья кость» [18]. Диаграмма наглядно отражает основные причинно-следственные связи предметной области и легко преобразуется в набор продукций. И диаграммы, и БЗ потом можно дополнять, развивая систему. Их использование здесь позволило визуализировать и систематизировать БЗ, облегчив её наполнение.

Для того, чтобы при помощи диаграммы Исикавы сформировать БЗ, необходимо сгруппировать транспортные происшествия на железной дороге. В [8] причины схода были разбиты на 5 групп, представленных на рисунке 1. Это:

- путь;
- вагоны;
- станция;
- локомотивная бригада;
- электрика.

Эти группы, в свою очередь, разбиваются на подгруппы для более детального анализа ситуации, что способствует наполнению БЗ новыми правилами. Для примера, разобьем на подгруппы группу причин «Путь». На рисунке 2 представлена относительно детализированная диаграмма «Сход состава с рельсов из-за неисправностей пути» [19].

Видно, что группа причин «Путь» разбивается на такие подгруппы, как [20-22]:

- шпалы (рисунок 3);
- изломы рельса (фрагмент ветви представлен на рисунке 4);
- стрелочный перевод и другие.

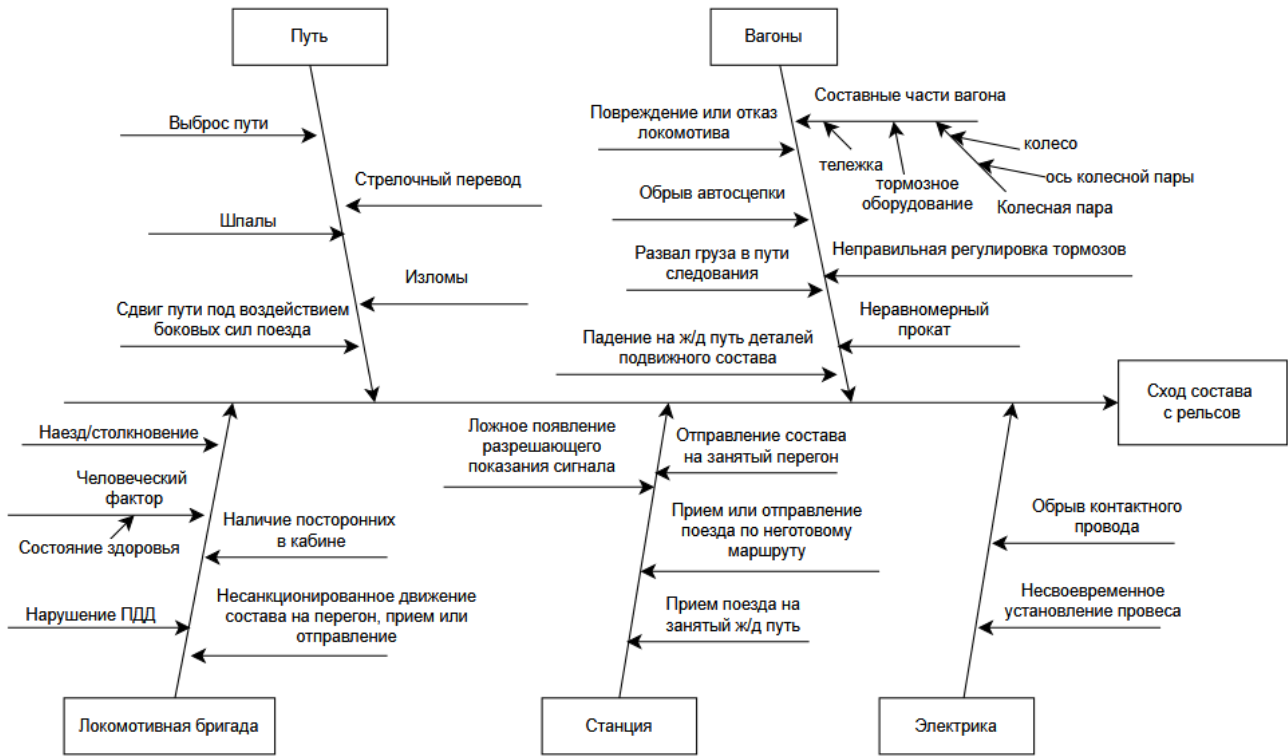


Рис.1. Диаграмма Исикавы «Причины схода состава с рельсов»

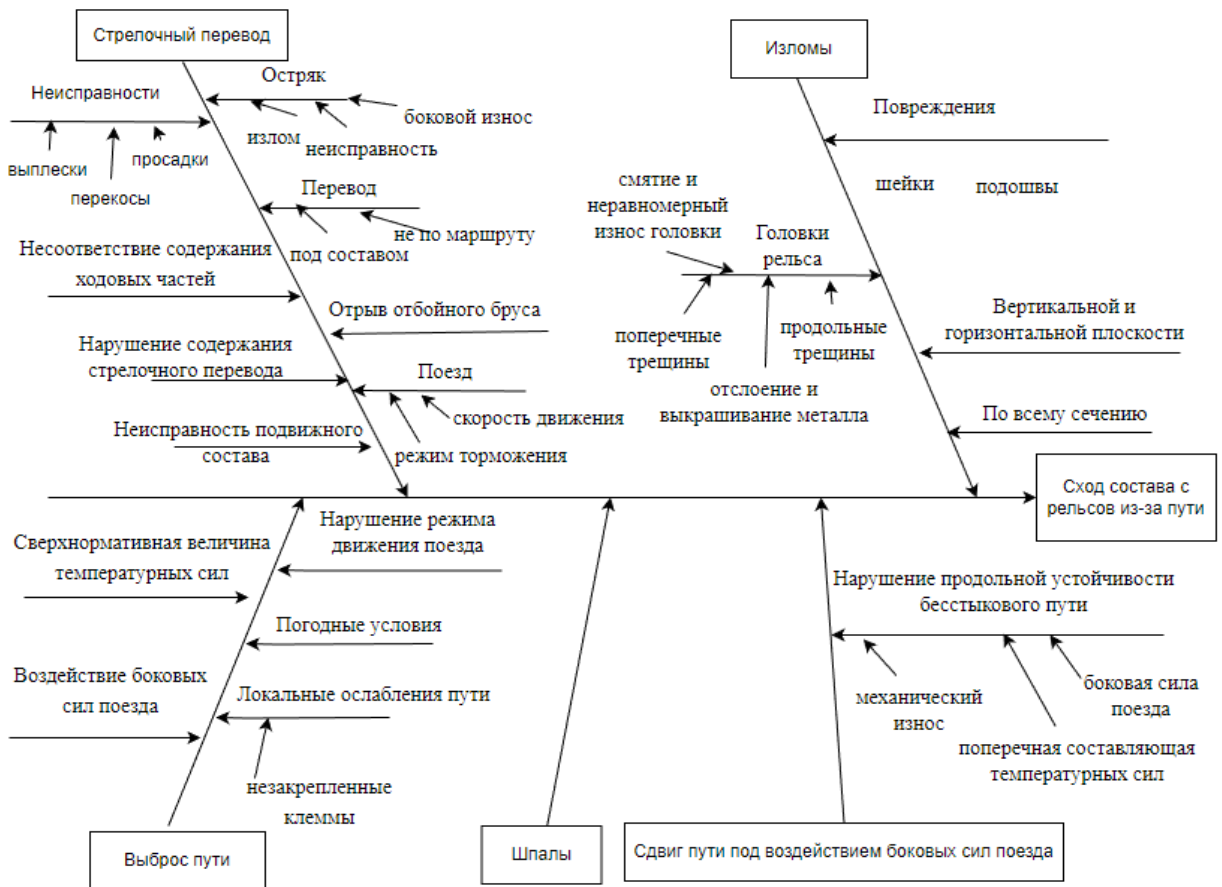


Рис. 2. Диаграмма Исикавы «Сход состава с рельсов из-за пути»

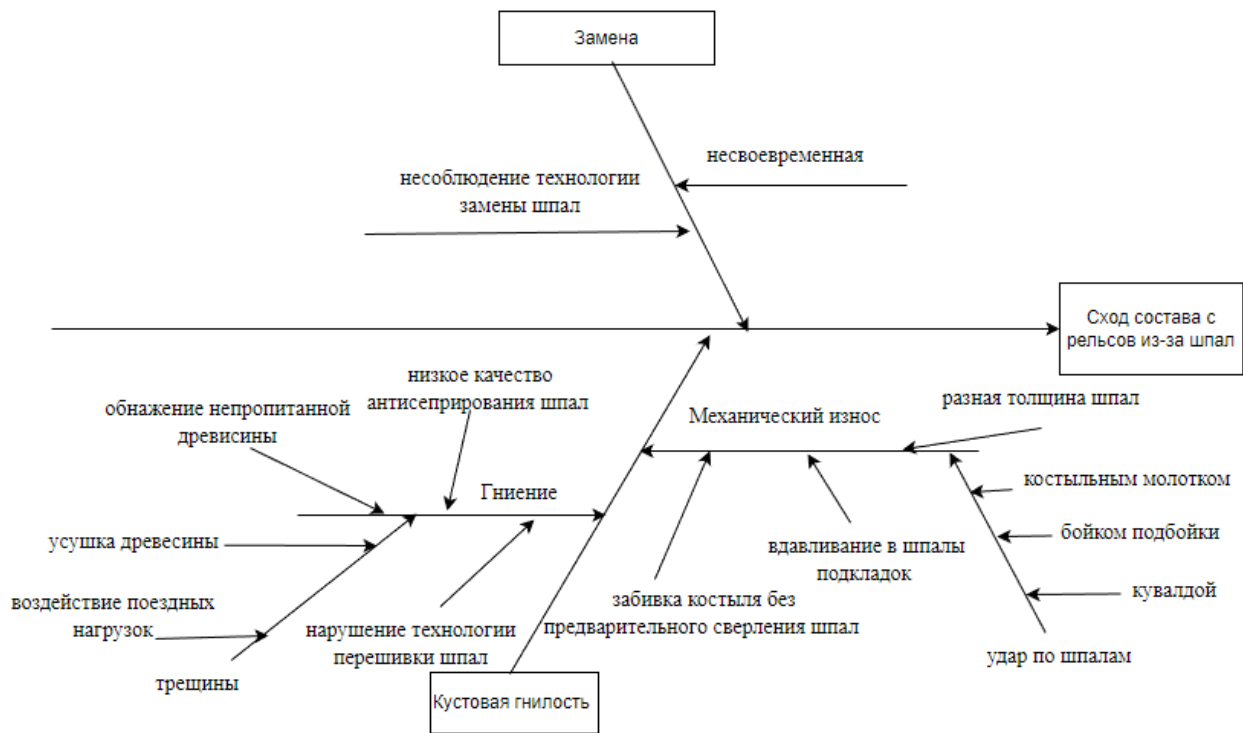


Рис. 3. Диаграмма Исикавы «Сход состава с рельсов из-за шпал»

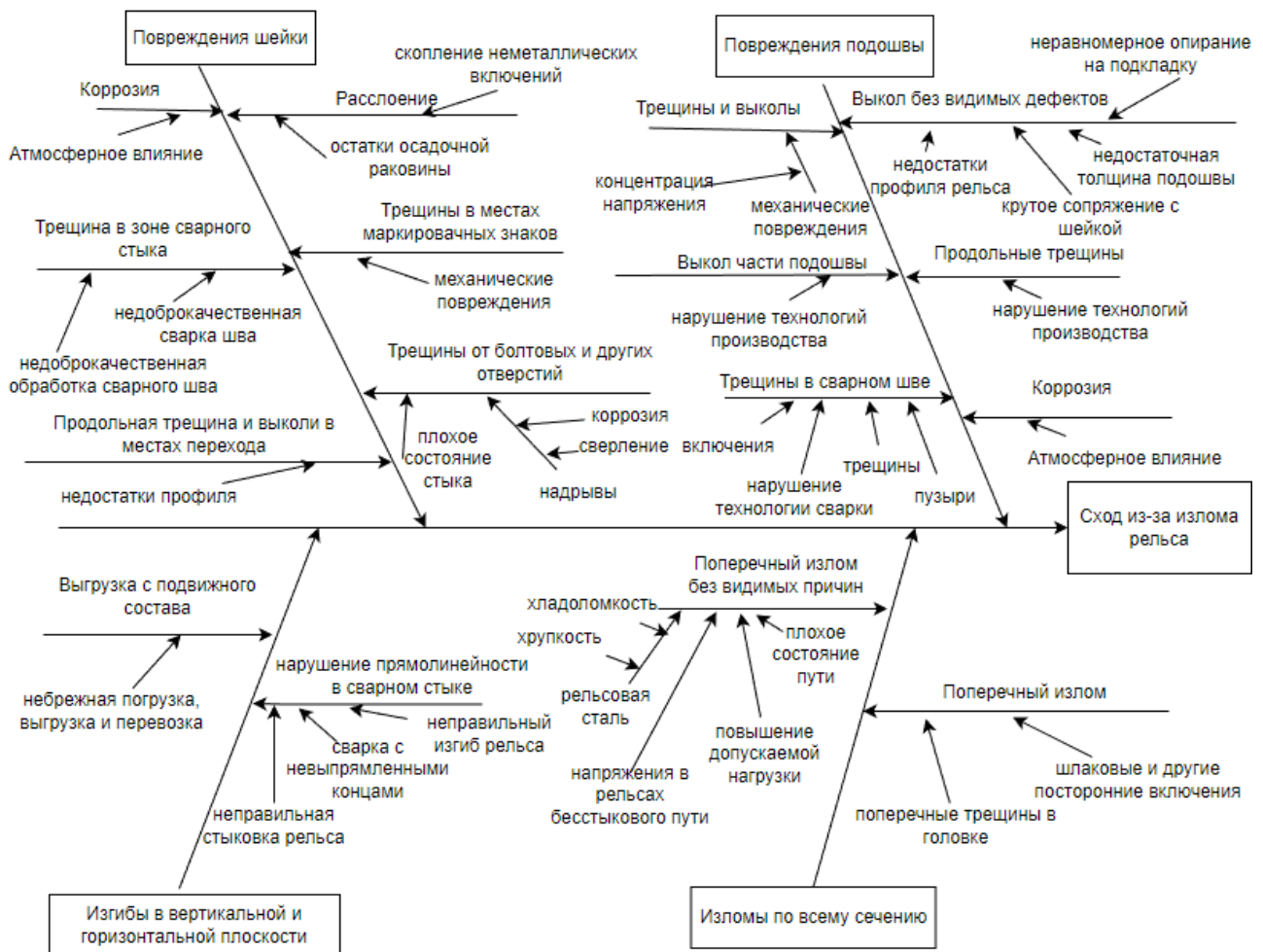


Рис. 4. Диаграмма Исикавы «Сход из-за излома рельса»

Аналогичным образом детализируются также такие причины, как «Сход состава с рельсов из-за неисправностей вагона», «Сход по вине локомотивной бригады», и так далее [19, 23, 24]. Например, группа причин «Вагоны» разбивается на такие подгруппы, как [25-29]:

- колесная пара;
- колесо;
- ось колесной пары;
- тележка вагона;
- кузов или рама полувагона;
- тормозное оборудование.

Источниками знаний для построения диаграмм выступают эксперты и документы. Например, для вагонов это РД 32 ЦВ 062-99 «Типовой технологический процесс работы пункта технической передачи вагонов в составе ПТО станции и контроля за сохранностью вагонного парка» [30] и 808-2017 ПКБ ЦВ «Инструкция по техническому обслуживанию вагонов в эксплуатации (Инструкция осмотрику вагонов)» [31]. В этих документах описаны организация работ по контролю за сохранностью вагонов предприятия, правила технического обслуживания вагонов, перечень неисправностей элементов вагона, а также приведены показатели контролируемых величин элементов вагона, показывающих, в каких пределах допустимы дефекты элементов вагона.

Пример диаграммы «Сход из-за колесной пары» представлен на рисунке 5.

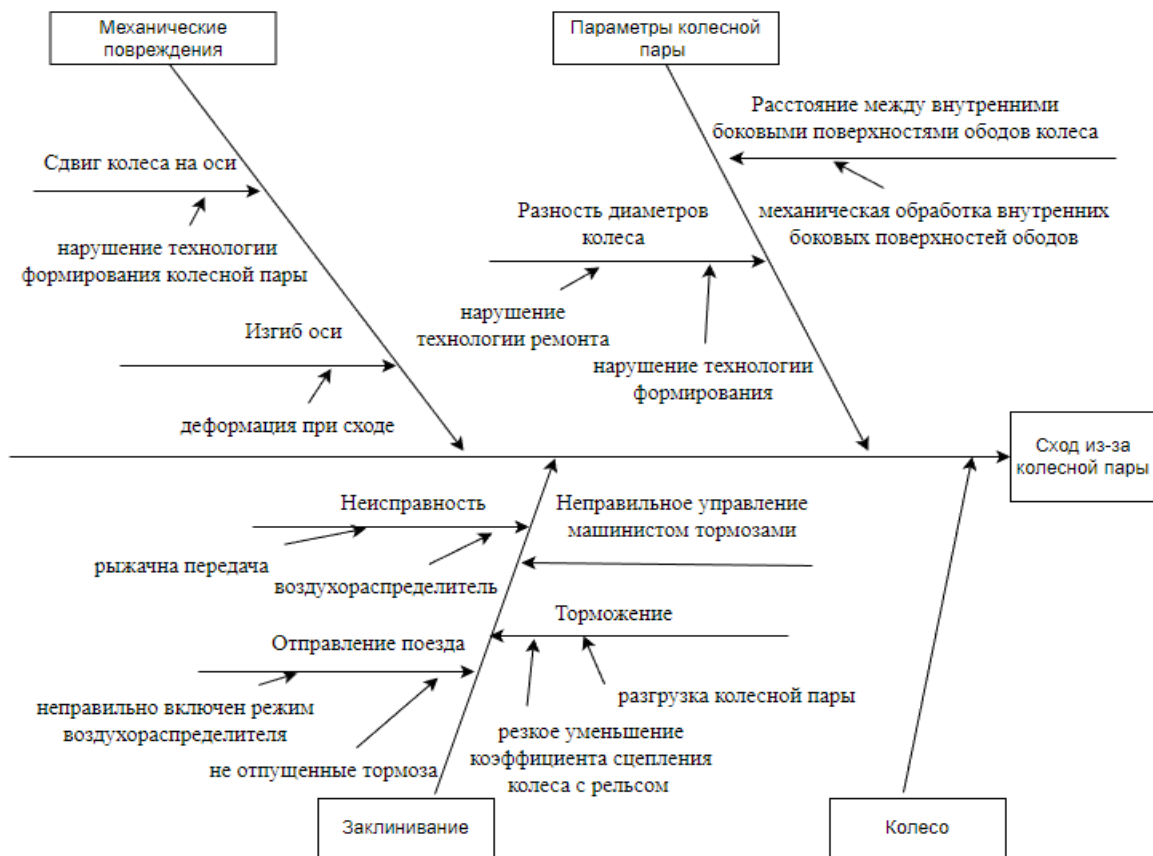


Рис. 5. Диаграмма Исикавы «Сход из-за колесной пары»

Хорошим вариантом для первоначального заполнения БЗ будет разбиение каждой группы причин как минимум на 3-5 подгруппы и формирование диаграммы для каждой подгруппы. Таким образом, получается до 5 групп причин, которые, в свою очередь, разбиваются на 4-6 подгрупп, каждая из которых содержит как минимум 3-4 категории факторов, состоящих из 2-6 факторов схода. В результате получаем минимальное число правил для БЗ, которых

будет достаточно для первоначального функционирования ЭС. Всего при разработке ЭС использовались восемнадцать диаграмм, каждая из которых при необходимости может быть развита и дополнена. Соответственно расширяется и БЗ. На данный момент ЭС содержит 684 факта и 431 правило.

4. Реализация БЗ ЭС. Для разработки прототипа ЭС и его БЗ была выбрана система моделирования правдоподобных рассуждений «Гераклит 2.Х», которая предназначена для разработки и эксплуатации декларативных баз знаний в различных предметных областях [32]. В системе используется продукционная модель знаний. Левая часть продукций (антецедент) может содержать скобочные выражения со связками И, ИЛИ, НЕ. Правая (консеквент) бесскобочная; допускает связки И, НЕ, либо ИЛИ, НЕ (связка ИЛИ по известным причинам в консеквенте нежелательна и в разрабатываемом прототипе не используется, однако инструментальная система такие связки обрабатывает). БЗ в «Гераклит 2.Х» реализована как реляционная база данных, состоящая из набора взаимосвязанных таблиц. Система таблиц включает:

- таблицу фактов (facts);
- таблицу узлов продукций (rulatt);
- таблицу частей ЕСЛИ продукций (rulif), содержащую ссылки на соответствующие факты и узлы продукций;
- таблицу частей ТО продукций (rulthen), содержащую ссылки на соответствующие факты и узлы продукций;
- таблицу частей ИНАЧЕ продукций (rulelse), содержащую ссылки на соответствующие факты и узлы продукций.

В системе реализуется присоединённый вывод: каждый шаг вывода сопровождается расчетом истинности заключения на основе истинностей посылок. Это позволяет, как в диалоге с пользователем, так и в диалоге с экспертом, обрабатывать ситуацию неполной уверенности. Кроме того, в системе используется векторная модель истинности (векторная семантика), обеспечивающая работоспособность машины вывода в условиях неполноты и противоречивости данных: входных фактов и правил. Причём компоненты вектора истинности могут иметь интервальные значения:

$$||a|| = \langle [a_{min}^+, a_{max}^+]; [a_{min}^-, a_{max}^-] \rangle, \quad (1)$$

где: $a_{min}^+, a_{max}^+, a_{min}^-, a_{max}^- \in [0,1]$ (для удобства шкала $[0,1]$ преобразована в целочисленную шкалу $0 \div 100$).

Подобное представление истинности позволяет, в частности, работать со связкой ИЛИ в правой части и выявлять противоречия в БЗ [16, 17, 32]. Частными случаями этой формализации являются строгое (классическое) и нечёткое (fuzzy) представления, что позволяет системе работать и с более привычными семантиками.

Продукции обрабатываются таким образом, чтобы обеспечить накопление истинности для антецедента продукции до того, как соответствующая продукция активизируется. Диалог при этом осуществляется за один проход БЗ (в этом смысле она похожа на миварные ЭС, описанные, к примеру, в [33-35]). Важным условием для этого является отсутствие в БЗ логических кругов.

Вывод в системе может быть пошаговый (гипотезы формируются по мере ответов на вопросы) и фронтальный (гипотезы формируются после ответов на все вопросы). Сама инструментальная система позволяет реализовывать прямой и обратный выводы, но по умолчанию предустановлен прямой пошаговый вывод. Для этого БЗ дополняется продукционными правилами выбора проверяемой гипотезы:

ЕСЛИ Проверяем гипотезу i , **ТО** Проверяемая гипотеза i .

Как следует из рисунка 1, гипотез здесь пять:

1. Возможная причина схода – путь (рисунки 2-4 и рисунок 6).
2. Возможная причина схода – вагоны.
3. Возможная причина схода – локомотивная бригада.
4. Возможная причина схода – станция.
5. Возможная причина схода – электрика.

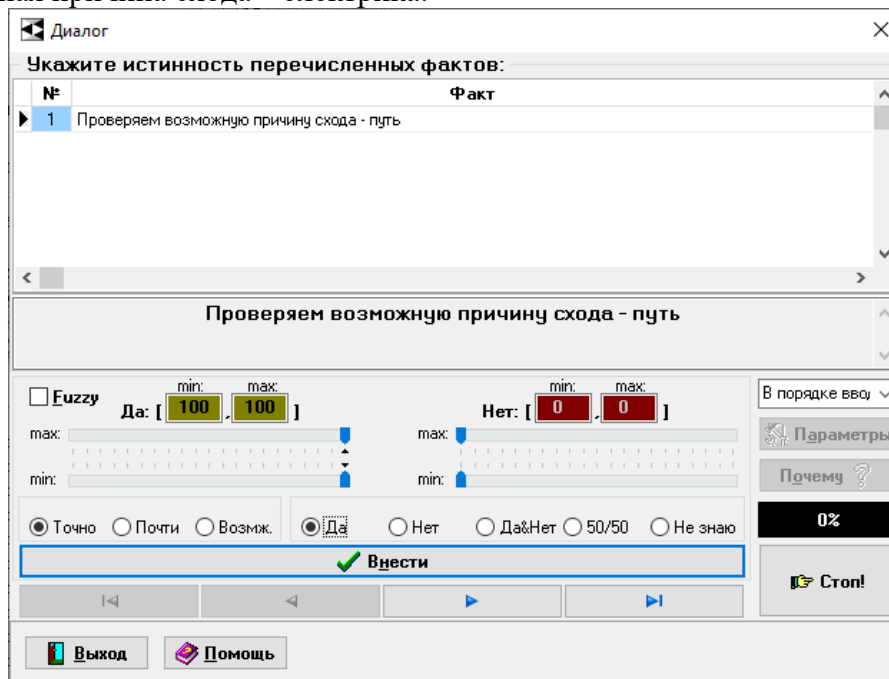


Рис. 6. Первоначальный вопрос для рассмотрения группы вопросов

Внутри каждой гипотезы факты группируются в ветви. К примеру, гипотеза «Путь» состоит из ветвей (рисунок 2):

1. Стрелочный перевод.
2. Изломы.
3. Выброс пути.
4. Шпалы.
5. Сдвиг пути под воздействием боковых сил поезда.

Пример начала диалога приведён на рисунке 6.

По каждой из гипотез также можно проверять не все ветви, а лишь те, которые требуют внимания. Достигается это с помощью правил вида:

ЕСЛИ Проверяемая гипотеза *i* **И** Проверяем ветвь *ij*, **ТО** Проверяемая ветвь *ij*.

Точно так же обрабатываются и все под-ветви, если они имеются. На нижнем уровне находятся правила вида:

ЕСЛИ Проверяемая ветвь *ij...k* **И** (Признак1 **ИЛИ** ... **ИЛИ** ПризнакN), **ТО** Результат.

Например, как в данном случае:

ЕСЛИ

Проверяем состояние стрелочного перевода
.И.(На стрелочном переводе наблюдаются выплески
.ИЛИ.На стрелочном переводе наблюдаются перекосы
.ИЛИ.На стрелочном переводе наблюдаются просадки)

ТО

Неисправен стрелочный перевод <[100,100];[0,0]>

Признаки здесь – это инструментально проверяемые обстоятельства схода. При необходимости признаки могут участвовать и в правилах верхних уровней.

Результатом диалога выступает гипотеза или любая из её ветвей/под-ветвей в зависимости от организации знаний. В рассматриваемой ЭС итоговыми заключениями выступают ветви (рисунок 7).

Следует обратить внимание, что высказывания, отвечающие невыбранным вариантам (гипотезам, ветвям/подветвям, признакам), в рамках модели (1) имеют интервальные значения истинности $||a|| = \langle [0,0]; [0,0] \rangle$ – «неопределённость» и в выводе не участвуют (как не участвуют в нём высказывания с истинностью $||a|| = \langle [0,0]; [1,1] \rangle$ – «строгая ложь») [32, 36].

Факты и правила в ЭС объединены в вычислительную сеть – двудольный граф – с двумя типами вершин: факты и продукции (гипотезы также рассматриваются как факты, но конечные, не входящие в левую часть ни одного из правил); вывод представляет собой вычисление на этой сети [32]. Это избавляет от необходимости использовать в машине вывода т.н. «рабочую память» для хранения активных фактов и правил и обеспечивает вывод за один проход БЗ (похожий приём используется и в упомянутом миварном подходе к разработке ЭС).

После того как пользователь ответит на все выбранные им вопросы, ЭС выведет полученные результаты (рисунок 7).

Выбор предпочтительной гипотезы осуществляется на основе лексикографического порядка: достоверность-определённость. Остальные показатели являются частью вывода системы и здесь носят вспомогательный характер.

Если выбранная версия не подтверждается, в окне с результатами диалога появляется соответствующее сообщение (рисунок 8). Подсистема объяснений («Почему?») предоставит ход рассуждений, приведших к результату.

Если на каких-то этапах пользователь неуверенно отвечал на вопросы или эксперт был не до конца уверен в тех или иных правилах, а также, возможно, поступали противоречивые сведения, ответ ЭС отразит это обстоятельство (рисунок 9).

Наконец, если эксперт хочет проверить другую гипотезу, рассуждения проводятся для неё аналогичным образом. Кроме того, как говорилось выше, можно проверять сразу несколько возможных гипотез и/или несколько ветвей в их составе; число задаваемых вопросов при этом возрастает.

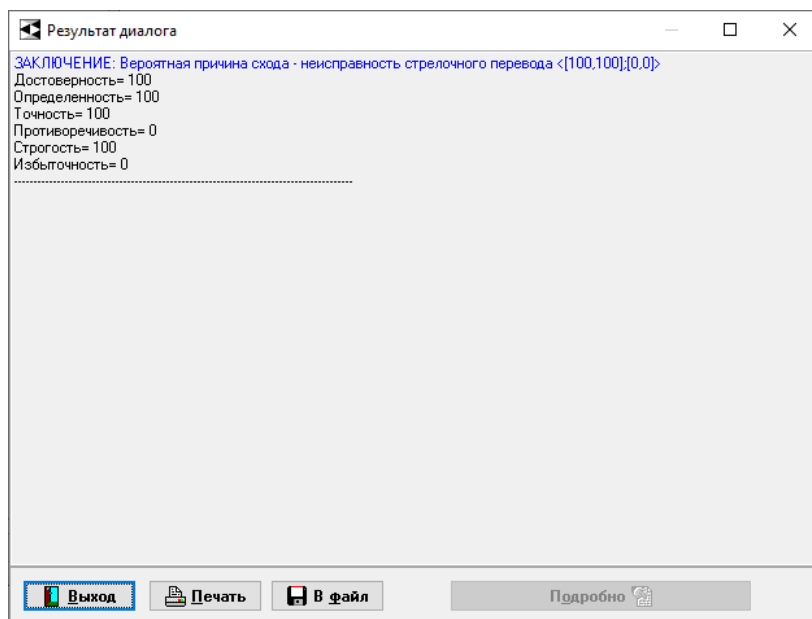


Рис. 7. Заключение системы при положительном исходе диалога по ветке «Возможная причина схода – путь»

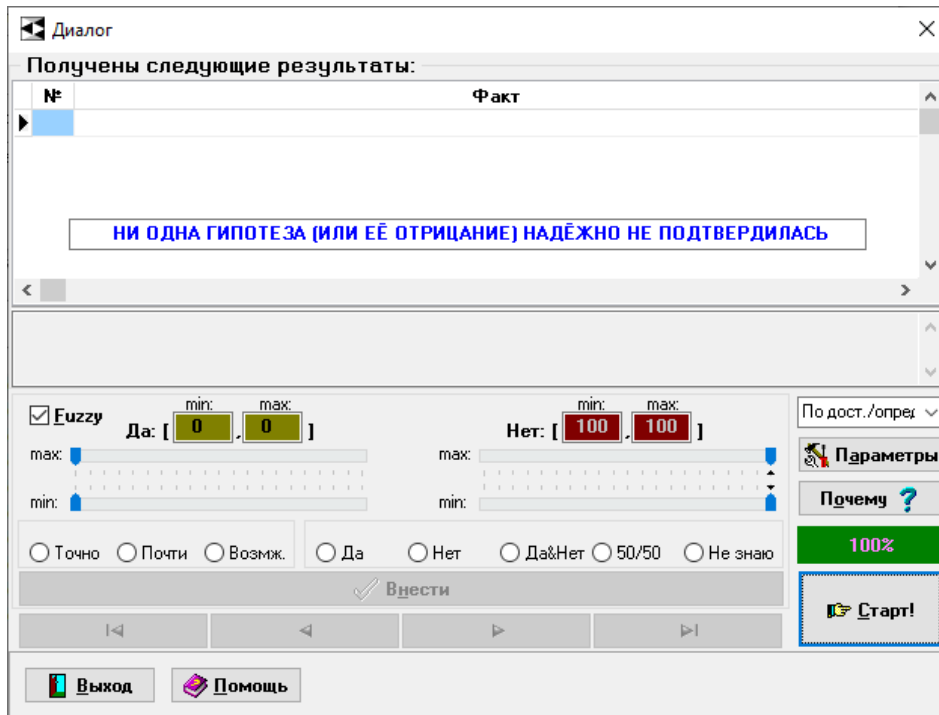


Рис. 8. Заключение системы при отсутствии подтверждающих фактов

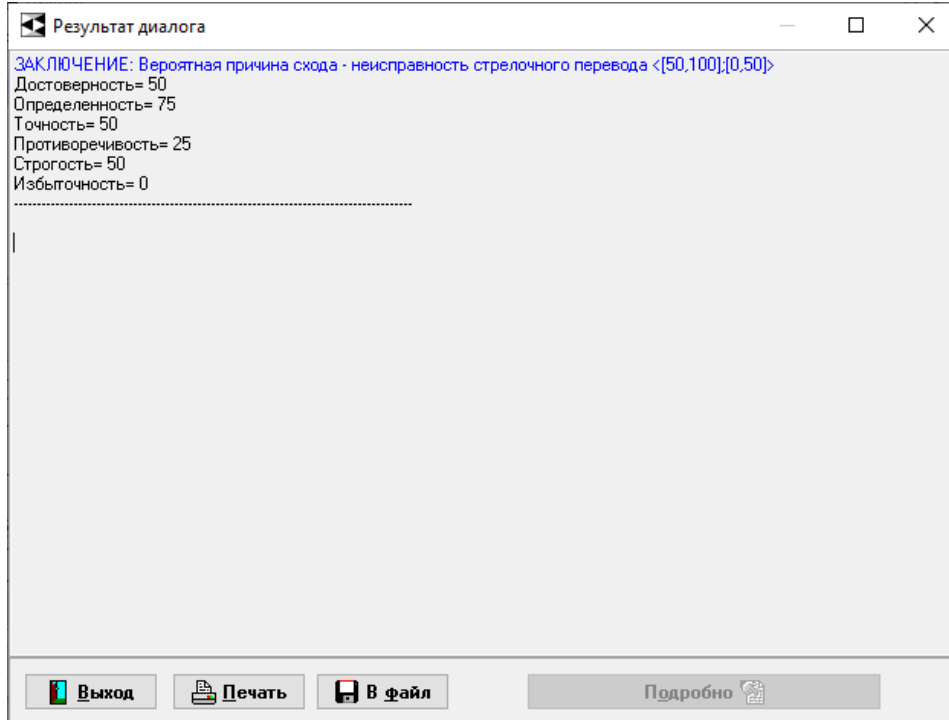


Рис. 9. Пример заключения при неуверенных ответах на вопросы или небесспорных правилах

Заключение. В работе рассмотрен один из подходов к разработке продукционной экспертной системы по анализу происшествий на железнодорожном транспорте. Обсуждаются этапы и особенности разработки её базы знаний и организации диалога. Прототип ЭС разработан только для анализа сходов, но принятая в ней (продукционная) модель знаний позволяет относительно легко расширить БЗ и на другие причины происшествий.

Система реализована в среде Гераклит 2.X, использующей вывод на основе логик с векторной семантикой, которые позволяют работать как с достоверной, так и с нечёткой, неопределённой и противоречивой информацией.

Список источников

1. Архангельская О. Обзор отрасли грузоперевозок в России / О. Архангельская, П. Галова, С. Трофимов, В. Махачев. Эрнст энд Янг – оценка и консультационные услуги, 2020. – 54 с.
2. Распоряжения ОАО «РЖД» от 26.01.2006 № 129р «О разработке порядка организации учета и формирования отчетности по нарушениям безопасности движения в поездной и маневровой работе, допущенным участниками перевозочного процесса на железных дорогах ОАО «РЖД». – URL: <https://jd-doc.ru/2006/yanvar-2006/12437-rasporyazhenie-oao-rzhd-ot-26-01-2006-n-129r/> (дата обращения: 07.02.2024).
3. Подобедова Е.В. Проблема разработки экспертной системы по анализу происшествий на железнодорожном транспорте / Е.В. Подобедова, М.С. Жукова // Молодая наука Сибири, 2022. – № 3(17). – С.184-191.
4. Технология использования экспертных систем. – URL: <http://opengl.org.ru/avtomatizirovannye-informatsionnye-tekhnologii-v-ekonomike/tekhnologiya-ispolzovaniya-ekspertnykh-sistem.html> (дата обращения: 07.02.2024).
5. Аршинский Л.В. Интеллектуальные информационные системы и технологии: учебное пособие / Л.В. Аршинский, М.С. Жукова. – Иркутск: ИрГУПС, 2023. – 128 с.
6. Головицына М. Информационные технологии в экономике. Лекция 11: Интеллектуальные задачи в экономике. Преимущества экспертных систем. – URL: <https://intuit.ru/studies/courses/3735/977/lecture/14689?page=8%23sect37> (дата обращения: 07.02.2024).
7. Приказ Министерства транспорта РФ от 18 декабря 2014 г. N 344 «Об утверждении Положения о классификации, порядке расследования и учета транспортных происшествий и иных событий, связанных с нарушением правил безопасности движения и эксплуатации железнодорожного транспорта» с изменениями от 29 июля 2016, 1 июня 2018. – URL: <https://base.garant.ru/70878628/> (дата обращения: 07.02.2024).
8. Нитежук М.С. Особенности информационной поддержки расследования происшествий на железнодорожном транспорте / М.С. Нитежук, Л.В. Аршинский // Информационные технологии и математическое моделирование в управлении сложными системами, 2021. – № 4(12). – С. 1-10. – DOI: 10.26731/2658-3704.2021.4(12).1-10.
9. Почему вагоны сходят с рельсов? – URL: <https://delta-trans.ru/news/113-html/>(дата обращения: 07.02.2024)
10. Распоряжение ОАО «РЖД» от 08.05.2015 № 1185р «Об утверждении Положения об организации расследования и учета транспортных происшествий и иных событий, связанных с нарушением правил безопасности движения и эксплуатации железнодорожного транспорта на инфраструктуре ОАО РЖД». – URL: <https://rulaws.ru/acts/Rasporyazhenie-OAO-RZHD-ot-08.05.2015-N-1185r/> (дата обращения: 07.02.2024).
11. Морозова В.А. Представление знаний в экспертных системах: учебное пособие / сост. В.А. Морозова, В.И. Паутов. – Екатеринбург: Изд-во Урал. ун-та, 2017. – 120 с.
12. Логунова Е.А. Обзор подходов к разрешению недостатков продукционной базы знаний системы логического вывода / Е.А. Логунова // Современные наукоемкие технологии, 2015. – № 9. – С. 46-48.
13. Смирнов В.В. Методы и средства верификации баз знаний в интегрированных экспертных системах: специальность 05.13.11 «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»: автореферат диссертации на соискание учёной степени кандидата технических наук / Смирнов Виталий Валерьевич; Московский инженерно-физический институт. – Москва, 2006. – 25 с.
14. Davis R. Applications of meta-level knowledge to the construction, maintenance, and use of large knowledge bases. Ph D. diss, Dept of computer science, Stanford univ.; 1976.
15. Nguyen T.A., Perkins W.A., Laffey J.T., Pecora D. Knowledge base verification. AI Magazine, 1987, vol. 8, no. 2, pp. 69-75.
16. Аршинский Л.В. Логика с векторной семантикой как средство верификации баз знаний / Л.В. Аршинский, А.А. Ермаков, М.С. Жукова // Онтология проектирования, 2019. – Т. 9. – №4 (34). – С. 510-521. DOI: 10.18287/2223-9537-2019-9-4-510-521.
17. Аршинский Л.В. Неклассические логики в задаче верификации продукционных баз знаний / Л.В. Аршинский, М.С. Жукова // Информационные и математические технологии в науке и управлении, 2020. – № 2(18). – С. 36-51. – DOI: 10.38028/ESI.2020.18.2.003.
18. Метод Исикавы / Е. Сушина. – URL: <https://busyspace.ru/articles/diagramma-isikavy> (дата обращения: 07.02.2024).
19. Почему разбиваются поезда. – URL: <http://eav.ru/publ1.php?publid=2003-09a09> (дата обращения: 07.02.2024).
20. Инструкция по содержанию деревянных шпал, переводных и мостовых брусьев железных дорог колеи 1520 мм. – URL: <https://epk-rzd.ru/wp-content/uploads/2015/09/CPT-410.pdf> (дата обращения: 07.02.2024).
21. Кусты негодных шпал. – URL: <http://helpiks.org/4-83426.html> (дата обращения: 07.02.2024).

22. Распоряжение ОАО «РЖД» от 23.10.2014 № 2499р «Об утверждении и введении в действие инструкции «Дефекты рельсов. Классификация, каталог и параметры дефектных и остродефектных рельсов». – URL: <https://cssrzd.ru/orders/2499.pdf> (дата обращения: 07.02.2024).
23. Распоряжение «Об утверждении Рекомендаций по предотвращению случаев обрывов автосцепных устройств грузовых вагонов» № 2260р от 02 ноября 2010. – URL: https://instructionsrzd.ucoz.ru/_ld/1/158_2260_.pdf (дата обращения: 07.02.2024).
24. С какими неисправностями запрещается выпускать локомотив в эксплуатацию. – URL: <https://helpiks.org/3-56794.html> (дата обращения: 07.02.2024).
25. Классификатор неисправностей вагонных колесных пар и их элементов 1.20.001-2007. – URL: <http://rcit.su/techinfo51.html> (дата обращения: 07.02.2024).
26. Неисправности кузовов вагонов, не допускаемые в эксплуатации. – URL: <https://gdzp.ru/poleznaya-informaciya/kuzova-gruzovyh-vagonov/neispravnosti-kuzovov-vagonov-ne-dopuskaemye-v-ekspluatácii/> (дата обращения: 07.02.2024).
27. Неисправности тормозного оборудования вагонов, с которыми запрещается ставить их в поезда. – URL: http://studopedia.ru/19_176078_neispravnosti-tormoznogo-oborudovaniya-vagonov-s-kotorimi-zapreshchaetsya-stavit-ih-v-poezda.html (дата обращения: 07.02.2024).
28. Основные неисправности и дефекты полувагона. – URL: <https://vagonmaster.com/information/remont-poluvaqonov.> (дата обращения: 07.02.2024).
29. Руководящий документ РД 32 ЦВ 052-2009 «Ремонт тележек грузовых вагонов». – URL: <https://www.regtransservice.ru/wp-content/uploads/2021/02/doc39.pdf> (дата обращения: 07.02.2024).
30. Руководящий документ РД 32 ЦВ 062-99 «Типовой технологический процесс работы пункта технической передачи вагонов в составе ПТО станции и контроля за сохранностью вагонного парка». – URL: <http://files.stroyinf.ru/Data2/1/4293835/4293835315.htm> (дата обращения: 07.02.2024).
31. Инструкция по техническому обслуживанию вагонов в эксплуатации (Инструкция осмотрищику вагонов) 808-2017 ПКБ ЦВ. – URL: <http://docs.cntd.ru/document/495876742> — Дата доступа 03.02.2023
32. 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.
33. 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.
34. 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
35. Чувиков Д.А. Об экспертной системе «Анализ ДТП», основанной на концепции миварного подхода / Д.А. Чувиков // Проблемы искусственного интеллекта, 2017 – № 2 (5). С. 78-88.
36. Аршинский Л.В. Интервальное оценивание истинности в системах автоматизированных рассуждений на основе VTF-логик / Л.В. Аршинский // Труды IV международной конференции «Идентификация систем и задачи управления». SICPRO'05. Москва 25-28 января 2005. – М.: ИПУ РАН, 2005. – С. 1061-1074.

Подобедова Елизавета Вячеславовна. Магистрант кафедры Информационные системы и защита информации Иркутского государственного университета путей сообщения, Author ID: 1233438, SPIN: 9296-4330, podobedova1998@mail.ru, Россия, Иркутск, Чернышевского, 15.

Жукова Марина Сергеевна. Старший преподаватель кафедры Информационные системы и защита информации Иркутского государственного университета путей сообщения, Author ID: 1007958, SPIN: 6502-4350, marino_@mail.ru, Россия, Иркутск, Чернышевского, 15.

Аршинский Леонид Вадимович. Д.т.н., доцент, профессор кафедры Информационные системы и защита информации Иркутского государственного университета путей сообщения, Author ID: 520252, SPIN: 9286-4084, ORCID: 0000-0001-5135-7921, larsh@mail.ru, Россия, Иркутск, Чернышевского, 15.

Expert system for the support of investigation of accidents in railway transport

Elizaveta V. Podobedova, Marina S. Zhukova, Leonid V. Arshinskiy

Irkutsk state transport university, Irkutsk, Russia, *larsh@mail.ru*

Abstract. The paper considers an approach to the development of an expert system for the analysis of accidents in railway transport. The stages and features of the development of its knowledge base, the possibilities of a tool environment and other issues are discussed. The prototype of the system was developed only for the analysis of derailments, but the knowledge model adopted in it makes it relatively easy to expand the system to other causes of accidents. The system is implemented in an environment using inference based on logic with vector semantics, which allows you to work with both reliable and fuzzy, uncertain and contradictory information.

Keywords: expert system, vector semantics, accident investigation, railway transport

References

1. Arkhangel'skaya O., Galova P., Trofimov S., Makhachev V. Obzor otrasli gruzoperevozok v Rossii [Overview of the cargo transportation industry in Russia 2020]. Ernst end Yang – otsenka i konsul'tatsionnyye uslugi [Ernst & Young - Valuation and advisory services], 2020, 54 p.
2. Rasporyazheniya OAO «RZHD» ot 26.01.2006 № 129r «O razrabotke poryadka organizatsii ucheta i formirovaniya otchetnosti po narusheniyam bezopasnosti dvizheniya v poyezdnoy i manevrovoy rabote, dopushchennym uchastnikami perevozochnogo protsessa na zheleznykh dorogakh OAO «RZHD» [Order of JSC Russian Railways dated January 26, 2006 No. 129r “On the development of a procedure for organizing accounting and reporting on traffic safety violations in train and shunting operations committed by participants in the transportation process on the railways of JSC Russian Railways”], available at: <https://jd-doc.ru/2006/yanvar-2006/12437-rasporyazhenie-oao-rzhd-ot-26-01-2006-n-129r/> (accessed: 02/07/2024).
3. Podobedova E.V., Zhukova M.S. Problema razrabotki ekspertnoy sistemy po analizu proisshestviy na zheleznodorozhnom transporte [The problem of developing an expert system for the analysis of accidents in railway transport]. Molodaya nauka Sibiri [Young science of Siberia], 2022, no. 3(17), pp. 184-191.
4. Tekhnologiya ispol'zovaniya ekspertnykh sistem [Technology of using expert systems], available at: <http://opengl.org.ru/avtomatizirovannyye-informatsionnyye-tekhnologii-v-ekonomike/tekhnologiya-ispolzovaniya-ekspertnykh-sistem.html> (accessed: 02/07/2024).
5. Arshinskiy L.V., Zhukova M.S. Intellektual'nyye informatsionnyye sistemy i tekhnologii: uchebnoye posobiye [Intelligent information systems and technologies: textbook]. Irkutsk, Irkutsk State Transport University, 2023, 128 p.
6. Golovitsyna M. Informatsionnyye tekhnologii v ekonomike. Lektsiya 11: Intellektual'nyye zadachi v ekonomike. Preimushchestva ekspertnykh sistem [Golovitsyna M. Information technologies in economics. Lecture 11: Intellectual tasks in economics. Advantages of expert systems]. Available at: <https://intuit.ru/studies/courses/3735/977/lecture/14689?page=8%23sect37> (accessed: 02/07/2024).
7. Prikaz Ministerstva transporta RF ot 18 dekabrya 2014 g. N 344 «Ob utverzhdenii Polozheniya o klassifikatsii, poryadke rassledovaniya i ucheta transportnykh proisshestviy i inykh sobytiy, svyazannykh s narusheniem pravil bezopasnosti dvizheniya i ekspluatatsii zheleznodorozhnogo transporta» s izmeneniyami ot 29 iyulya 2016, 1 iyunya 2018 [Order of the Ministry of Transport of the Russian Federation dated December 18, 2014 N 344 “On approval of the Regulations on the classification, procedure for investigation and recording of transport accidents and other events related to violation of traffic safety rules and operation of railway transport” as amended on July 29, 2016, June 1, 2018]. Available at: <https://base.garant.ru/70878628/> (accessed: 02/07/2024).
8. Nitezhuik M.S., Arshinskiy L.V. Osobennosti informatsionnoy podderzhki rassledovaniya proisshestviy na zheleznodorozhnom transporte [Features of information support for the investigation of accidents in railway transport]. Informatsionnyye tekhnologii i matematicheskoye modelirovaniye v upravlenii slozhnymi sistemami [Information technologies and mathematical modeling in the management of complex systems], 2021, no. 4(12), pp. 1-10, DOI: 10.26731/2658-3704.2021.4(12).1-10.
9. Pochemu vagony skhodyat s rel'sov? [Why do carriages go off the rails?]. Available at: <https://delta-trans.ru/news/113-html/> (accessed: 02/07/2024)
10. Rasporyazheniye OAO «RZHD» ot 08.05.2015 № 1185r «Ob utverzhdenii Polozheniya ob organizatsii rassledovaniya i ucheta transportnykh proisshestviy i inykh sobytiy, svyazannykh s narusheniem pravil bezopasnosti dvizheniya i ekspluatatsii zheleznodorozhnogo transporta na infrastrukture OAO RZHD» [Order of JSC Russian Railways dated 05/08/2015 No. 1185r “On approval of the Regulations on the organization of investigation and recording of transport accidents and other events related to violation of traffic safety rules and operation of railway

- transport on the infrastructure of JSC Russian Railways”]. Available at: <https://rulaws.ru/acts/Rasporyazhenie-OAO-RZHD-ot-08.05.2015-N-1185r/> (accessed: 02/07/2024).
11. Morozova V.A., Pautov V.I. Predstavleniye znaniy v ekspertnykh sistemakh : uchebnoye posobiye [Knowledge representation in expert systems: textbook]. Yekaterinburg, Ural State University, 2017, 120 p.
 12. Logunova Ye.A. Obzor podkhodov k razresheniyu nedostatkov produktsionnoy bazy znaniy sistemy logicheskogo vyvoda [Review of approaches to resolving the shortcomings of the production knowledge base of a logical inference system]. Modern high technologies, 2015, no. 9, pp. 46-48.
 13. Smirnov V.V. Metody i sredstva verifikatsii baz znaniy v integrirovannykh ekspertnykh sistemakh [Methods and means of verifying knowledge bases in integrated expert systems]. Abstract of Ph.D thesis, Moscow, 2006, 25 p.
 14. Davis R. Applications of meta-level knowledge to the construction, maintenance, and use of large knowledge bases. Ph D. diss., Dept of computer science, Stanford univ., 1976.
 15. Nguyen T.A., Perkins W.A., Laffey J.T., Pecora D. Knowledge base verification. AI Magazine, 1987, vol. 8, no. 2, pp. 69-75.
 16. Arshinskiy L.V., Yermakov A.A., Nitezhuik M.S. Logiki s vektornoy semantikoy kak sredstvo verifikatsii baz znaniy [Logic with vector semantic as a means of knowledge bases verification]. Ontologiya dizayna [Ontology of designing], 2019, vol. 9, no. 4 (34), pp. 510-521, DOI: 10.18287/2223-9537-2019-9-4-510-521.
 17. Arshinskiy L.V., Zhukova M.S. Neklassicheskiye logiki v zadache verifikatsii produktsionnykh baz znaniy [Non-classical logics in rule-based system verification problem]. Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii [Information and mathematical technologies in science and management], 2020, no. 2(18), pp. 36-51, DOI: 10.38028/ESI.2020.18.2.003.
 18. Sushinina E. Metod Isikavy.[Ishikawa method]. Available at: <https://busyspace.ru/articles/diagramma-isikavy> (accessed: 02/07/2024).
 19. Pochemu razbivayutsya poyezda [Why do trains crash]. Available at: <http://eav.ru/publ1.php?publid=2003-09a09> (accessed: 02/07/2024).
 20. Instruktsiya po sodержaniyu derevyannykh shpal, perevodnykh i mostovykh brus'yev zheleznykh dorog kolei 1520 mm. [Instructions for the maintenance of wooden sleepers, transfer beams and bridge beams for 1520 mm gauge railways]. Available at: <https://epk-rzd.ru/wp-content/uploads/2015/09/CPT-410.pdf> (accessed: 02/07/2024).
 21. Kusty negodnykh shpal [Unusable sleepers]. Available at: <http://helpiks.org/4-83426.html> (accessed: 02/07/2024).
 22. Rasporyazheniye OAO «RZHD» ot 23.10.2014 № 2499r Ob utverzhenii i vvedenii v deystviye instruktsii "Defekty rel'sov. Klassifikatsiya, katalog i parametry defektnykh i ostrodefektnykh rel'sov" [Order of JSC "Russian Railways" dated October 23, 2014 No. 2499r On the approval and implementation of the instruction "Rail Defects. Classification, catalog and parameters of defective and severely defective rails"]. Available at: <https://cssrdz.ru/orders/2499.pdf> (accessed: 02/07/2024).
 23. Rasporyazheniye “Ob utverzhenii Rekomendatsiy po predotvrashcheniyu sluchayev obryvov avtostsepykh ustroystv gruzovykh vagonov” № 2260r ot 02 noyabrya 2010 [Order “On approval of Recommendations for preventing cases of breaks in automatic coupling devices of freight cars” No. 2260r dated November 2, 2010]. Available at: https://instructionsrdz.ucoz.ru/_ld/1/158_2260_..pdf (accessed: 02/07/2024).
 24. S kakimi neispravnostyami zapreshchayetsya vypuskat' lokomotiv v ekspluatatsiyu [What defects are prohibited from putting a locomotive into service]. Available at: <https://helpiks.org/3-56794.html> (accessed: 02/07/2024).
 25. Klassifikator neispravnostey vagonnykh kolesnykh par i ikh elementov 1.20.001-2007 [Classifier of faults in carriage wheel sets and their elements 1.20.001-2007]. Available at: <http://rcit.su/techinfo51.html> (accessed: 02/07/2024).
 26. Neispravnosti kuzovov vagonov, ne dopuskayemyye v ekspluatatsii [Malfunctions of car bodies that are not allowed in operation], Available at: <https://gdzp.ru/poleznaya-informaciya/kuzova-gruzovykh-vagonov/neispravnosti-kuzovov-vagonov-ne-dopuskaemye-v-ekspluatatsii/> (accessed: 02/07/2024).
 27. Neispravnosti tormoznogo oborudovaniya vagonov, s kotorymi zapreshchayetsya stavit' ikh v poyezda [Malfunctions of the brake equipment of car with which it is prohibited to place them on trains]. Available at: http://studopedia.ru/19_176078_neispravnosti-tormoznogo-oborudovaniya-vagonov-s-kotorimi-zapreshchaetsya-stavit-ih-v-poezda.html (accessed: 02/07/2024).
 28. Osnovnyye neispravnosti i defekty poluvagona [Main faults and defects of the gondola car]. Available at: <https://vagonmaster.com/information/remont-poluvagonov>. (accessed: 02/07/2024).
 29. Rukovodyashchiy dokument RD 32 TSV 052-2009 “Remont telezhek gruzovykh vagonov” [Guiding document RD 32 TsV 052-2009 “Repair of freight car bogies”]. Available at: <https://www.regtransservice.ru/wp-content/uploads/2021/02/doc39.pdf> (accessed: 02/07/2024).

30. Rukovodyashchiy dokument RD 32 TSV 062-99 “Tipovoy tekhnologicheskiy protsess raboty punkta tekhnicheskoy peredachi vagonov v sostave PTO stantsii i kontrolya za sokhrannost'yu vagonnogo parka” [Guiding document RD 32 TsV 062-99 “Typical technological process of operation of the point of technical transfer of cars as part of the station’s technical maintenance department and monitoring the safety of the car fleet”]. Available at: <http://files.stroyinf.ru/Data2/1/4293835/4293835315.htm> (accessed: 02/07/2024).
31. Instruktsiya po tekhnicheskomu obsluzhivaniyu vagonov v ekspluatatsii (Instruktsiya osmotrshchiku vagonov) 808-2017 PKB TSV [Instructions for the maintenance of wagons in operation (Instructions for wagon inspectors) 808-2017]. Available at: <http://docs.cntd.ru/document/495876742> (accessed: 02/07/2024).
32. Arshinskiy L.V. Osobennosti raboty mashiny vyvoda sistemy modelirovaniya pravdopodobnykh rassuzhdeniy «Geraklit» [Features of the operation of the inference machine 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.
33. Varlamov O.O. Evolyucionnyye 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]. Moscow, Radio i svyaz' [Radio and communication], 2002, 288 p., EDN RWTCOP.
34. Varlamov O.O. Sozdaniye Bol'shikh Znaniy i rasshireniye oblastey primeneniya mivarnykh tekhnologiy logicheskogo iskusstvennogo intellekta [Creating Big Knowledge and expanding the applications of mivar technologies of logical artificial intelligence]. *Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii* [Information and mathematical technologies in science and management], 2023, no. 4(32), pp. 30-41, DOI: 10.25729/ESI.2023.32.4.003.
35. Chuvikov D.A. Ob ekspertnoy sisteme «Analiz DTP», osnovannoy na kontseptsii mivarnogo podkhoda [About the expert system “Analysis MVA”, based on the concept of the mivar approach]. *Problems of artificial intelligence*, 2017, no. 2 (5), pp. 78-88.
36. Arshinskiy, L.V. Interval'noe ocenivanie istinnosti v sistemah avtomatizirovannykh rassuzhdenij na osnove VTF-logik [Interval estimation of the truth in the systems of automated reasoning based on the VTF-logics]. *Trudy IV mezhdunarodnoj konferencii «Identifikacija sistem i zadachi upravleniya»*. SICPRO'05. Moskva 25-28 janvarja 2005 [Proc. of IV International conference “System Identification and Control Problems”], Moscow, Institute of Control Sciences of RAS, 2005, pp. 1061-1074.

Podobedova Elizaveta Vyacheslavovna. Master's student of the department information systems and information security of the Irkutsk state transport university, Author ID: 1233438, SPIN: 9296-4330, podobedova1998@mail.ru.

Zhukova Marina Sergeevna. Senior lecturer of the department information systems and information security of Irkutsk state transport university, Author ID: 1007958, SPIN: 6502-4350, marino_@mail.ru.

Arshinskiy Leonid Vadimovich. Doctor of technical sciences, associate professor, professor of the department information systems and information security of Irkutsk state transport university, Author ID: 520252, SPIN: 9286-4084, ORCID: 0000-0001-5135-7921, larsh@mail.ru.

Статья поступила в редакцию 22.02.2024; одобрена после рецензирования 07.05.2024; принята к публикации 08.10.2024.

The article was submitted 02/22/2024; approved after reviewing 05/07/2024; accepted for publication 10/08/2024.

Математические, информационные и интеллектуальные технологии в энергетике

УДК 620.91

DOI:10.25729/ESI.2024.35.3.006

Первичный анализ работы гибридной микросети лабораторного масштаба с использованием системы мониторинга

Баденко Владислав Вадимович, Козлов Александр Николаевич

Институт систем энергетики им. Л.А. Мелентьева СО РАН,

Россия, Иркутск, *badenko@isem.irk.ru*

Аннотация. Одним из направлений современной энергетической промышленности является разработка систем, включающих в себя использование солнечной энергии, энергии ветра и переработку вторичных энергоресурсов (отходов). В работе рассматривается энергетическая система, обозначаемая как гибридная микросеть лабораторного масштаба, включающая в себя использование таких источников. Энергосистема включает в себя солнечные панели, газификатор биомассы и электрогенератор. Большую значимость для таких систем представляют инструменты мониторинга для анализа режимов работы и внесения модификаций. Исходя из этого, была разработана система мониторинга для отслеживания различных параметров лабораторного оборудования. Благодаря разработанной системе были получены данные о работе микросети, выявлены ее недостатки и узкие места.

Ключевые слова: солнечная энергия, биомасса, мониторинг

Цитирование: Баденко В.В. Первичный анализ работы гибридной микросети лабораторного масштаба посредством системы мониторинга / В.В. Баденко, А.Н. Козлов // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 73-81. – DOI: 10.25729/ESI.2024.35.3.006.

Введение. Энергетические системы, в основе которых лежит использование возобновляемых источников энергии, вызывают немалый интерес в исследовательском сообществе [1]. Одним из направлений современной энергетической промышленности является разработка систем, основанных на использовании солнечной энергии, энергии ветра [2] и переработке вторичных энергоресурсов (отходов) [3]. Такие системы особенно актуальны для развивающихся стран и районов, отрезанных от подключения к крупным энергосетям [4-5]. В энергосистему также может быть заложена генерация электроэнергии из невозобновляемых источников (горючих топлив), как в виде резервной альтернативы [6], так и в виде одного из основных узлов. В последнем случае комбинация источников используется для уменьшения расхода топлива и выбросов.

Гибридные энергосистемы местного масштаба часто описываются понятием «микросеть». Микросеть – автономная система электроснабжения с мощностью от нескольких кВт до нескольких мВт [7]. В процессе создания микросети необходимо учитывать разнообразные ограничения [8], возникающие при согласовании узлов микросети. Многие возникающие задачи не могут быть решены одновременно и требуют применения многокритериального подхода и поиска компромиссных решений [9]. В общем случае, мультиэнергетическая микросеть должна быть экономически эффективной, обеспечивать минимальное воздействие на окружающую среду, учитывать требования к качеству электроэнергии, а также иметь высокий уровень по другим показателям эффективности, например, устойчивости [10].

Идея применения гибридных энергосистем, объединяющих в себе различные источники энергии, базируется на взаимной компенсации недостатков. Например, комбинированное применение стандартного электрогенератора на горючем топливе в темное время суток и солнечных панелей в светлое позволяет повысить устойчивость энергоснабжения при экономии ресурсов [11]. Важной особенностью таких систем является модульность и распределенная генерация. Преимуществом микросетей является возможность интеграции в существующие

энергетические сети, поскольку микросеть может быть создана в различных конфигурациях в зависимости от местных условий окружающей среды.

Эксплуатация и разработка гибридных установок требуют наличия систем управления, подстраивающихся под запросы потребителей и прогнозирующих различные возмущения, а также систем хранения энергии и мониторинга. Это особенно актуально для установок, основанных на солнечной энергии, в силу зависимости от времени суток. Из-за различных энергоисточников и накопителей требуется рассмотрение характерных режимов для учета возможных узких мест при формировании структуры мультиэнергетической микросети, т.е. состава и установленной мощности оборудования [12].

Анализ работы микросети строится на мониторинге функционирования всех узлов в различных условиях. Исследование поведения микросети особенно важно при ее разработке и тестировании [13]. На основании собранных данных можно осуществлять модернизацию узлов микросети, выбирать режимы работы и устанавливать ограничивающие условия. Размер выборки для анализа имеет прямое влияние на качество анализа. Частота обновления данных также играет большую роль. Лучшим решением данной проблемы является система мониторинга с использованием глобальной сети интернет. Отслеживание характеристик системы с достаточно высокой частотой позволяет принимать своевременные решения о внесении изменений в конкретный узел сети. Такой подход в перспективе можно использовать для осуществления удаленного управления, в том числе и для аварийной остановки. Применение информационных технологий при анализе работы гибридных сетей также упрощает привлечение различных специалистов к процессам модернизации.

2. Материалы и методы. Для исследований была создана установка (рисунок 1), сочетающая в себе солнечные панели, установку по газификации древесной биомассы, электрогенератор, функционирующий на синтез-газе, систему хранения электроэнергии посредством аккумуляторов, систему управления, систему мониторинга.

Компоненты микросети выбирались с учетом их доступности и стоимости при обеспечении необходимых требований по функционированию. Также использовалось оборудование, имеющееся в лаборатории.

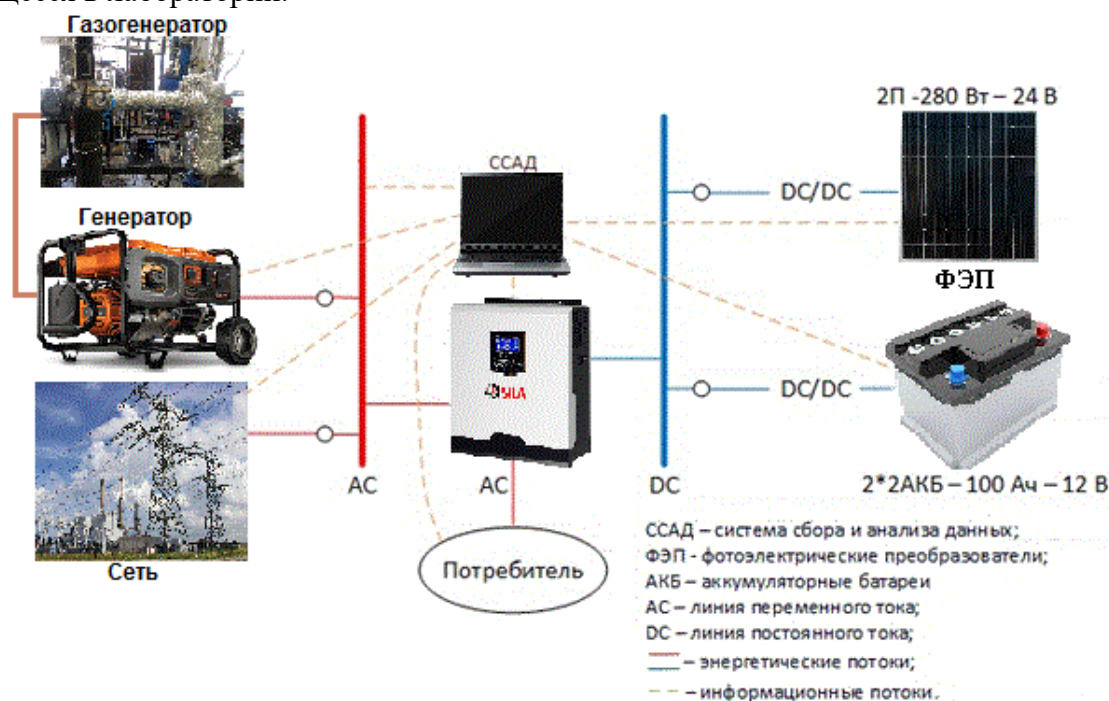


Рис. 1. Схема микросети

Установка по газификации биомассы состоит из реактора, системы нагрева, линии фильтрации смолы, системы подачи газа. Исходное сырье закладывается в реакционную зону на

колосниковую решетку, затем осуществляется подача инертного газа или окислителя. Газообразные продукты реакции проходят через многоступенчатую систему фильтрации и затем подаются в следующий узел для генерации энергии.

Для энергетического использования продуктов термохимической конверсии биомассы был модифицирован бытовой бензиновый электрогенератор СПЕЦ-HG-7500. Для адаптации устройства под синтез-газ были заменены системы впуска и зажигания.

Узел, основанный на получении электроэнергии из солнечных лучей, основывается на фотовольтаических элементах SilaSolar SIM280 (5ВВ) общей мощностью 560 Вт, установленных на открытом воздухе. Панели подключены к гибриднему солнечному инвертору SILA V 1000P – это многофункциональный инвертор/зарядное устройство мощностью 1 кВт (кратковременно 2 кВт), сочетающий в себе функции инвертора, контроллера заряда солнечных батарей и зарядного устройства от сети 220В для обеспечения бесперебойного электроснабжения, с возможностью выбора приоритетов заряда и нагрузки.

Управление и координация источников энергии осуществляется при помощи ПО Watchpower [14] (рисунок 2). Программа позволяет осуществлять управление микросетью и фиксировать показания измерительных приборов (мощность нагрузки, входное/выходное сетевое напряжение и частота, напряжение блока батарей, напряжение солнечных панелей и их коэффициент мощности). Также в программу передаются сообщения о неисправностях.

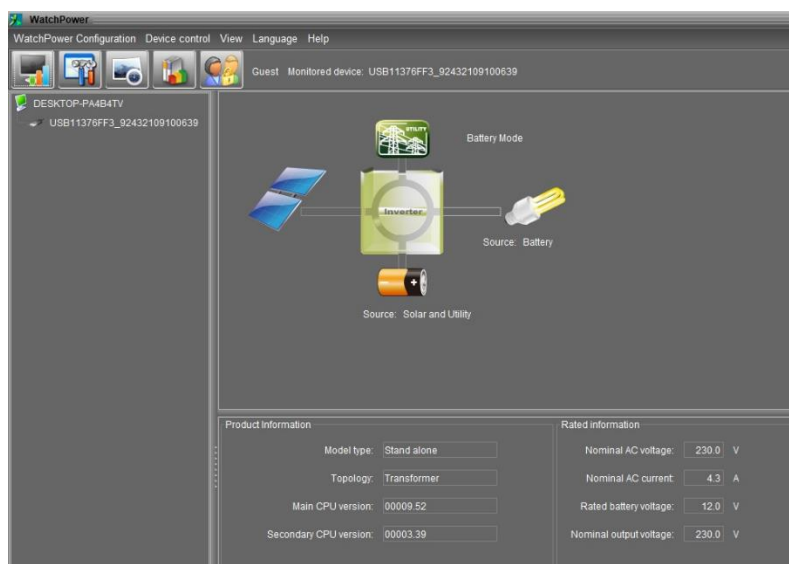


Рис. 2. Внешний вид интерфейса программы WatchPower

Хранение энергии осуществляется за счет четырех параллельно соединенных аккумуляторов типа AGM. Рабочее напряжение составляет 11-14 В, емкость блока аккумуляторов - 400 А•ч (емкость каждого аккумулятора - 100 А•ч).

Система передачи данных основывается на графическом представлении параметров микросети в качестве функции от времени. Концепция удаленного управления микросетью в данном исследовании не используется, поскольку требует отдельной системы обеспечения безопасности и не имеет определяющего значения для экспериментальной микросети, характерного для коммерческих установок, и будет описана в последующих работах. Графическое представление характеристик осуществляется на основе бесплатной графической библиотеки Google charts [15]. Передача данных может осуществляться как через открытый бесплатный репозиторий (например, GitHub), так и через серверную структуру организации ИСЭМ СО РАН. Страница с отображением графической информации основывается на языке гипертекстовой разметки HTML с применением CSS и JavaScript. Отслеживаемыми характеристиками микросети являются: операционные характеристики отдельного узла (датчики температуры,

системы отслеживания работоспособности), характеристики поставляемой в сеть электрической энергии (напряжение, частота, мощность), параметры текущего состояния системы хранения электроэнергии (емкость, напряжение, токи утечки).

3. Оценка работы микросети. Для оценки работы микросети необходимы контроль и анализ различных ее узлов. В данной работе основное внимание уделяется солнечным элементам и системе мониторинга.

3.1. Анализ работы солнечных батарей. Работа узла с фотовольтаическими элементами анализировалась в течение трех месяцев. Рассматривались следующие характеристики: изменение напряжения на аккумуляторах, изменение напряжения на солнечных панелях, изменение мощности нагрузки, фиксировалось использование внешней сети. Микросеть как подключалась к сети, так и отключалась для проведения эксперимента с автономной работой. В качестве потребителей для исследований были использованы система освещения и система вентиляции в служебном помещении. Текущее расположение панелей является фиксированным. Механизм адаптации под наиболее оптимальное для календарного периода или времени суток положение повысит эффективность данного узла микросети. Однако, для исследования это допущение не имеет критической важности.

3.2. Система мониторинга. Идея системы мониторинга основана на предоставлении удаленного доступа к данным работы микросети. Логическая схема взаимодействия с данными показана на рисунке 3.

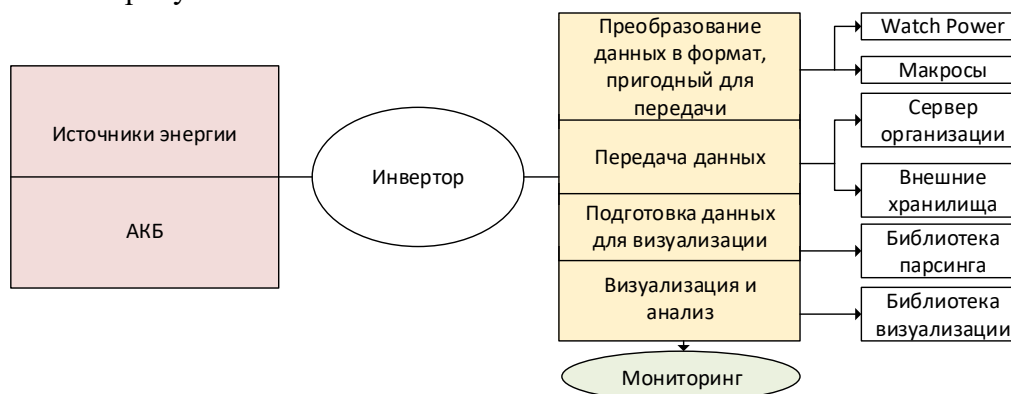


Рис. 3. Логическая схема взаимодействия с данными

Важным условием при проектировании микросети было объединение регистрируемых потоков данных в едином формате, пригодном для анализа и сопоставления. Это было реализовано корректным выбором инвертора и сопутствующего программного обеспечения, а также использованием наборов макросов. Регистрируемые данные содержатся в формате базы данных и конвертируются в пригодный для передачи формат CSV.

Передача данных реализована на текущий момент в виде статической ссылки на таблицу, которая может быть размещена как на базе доступного внешнего хранилища, так и на внутренних серверных хранилищах. Обновление данных происходит с суточной частотой, этого достаточно для функций мониторинга. В дальнейших работах будет осуществлен переход на обновление в режиме реального времени.

Обработка данных происходит на стороне серверной структуры, с использованием языка JavaScript и бесплатной библиотеки Papaparse [16]. В ходе обработки происходят необходимые преобразования формата данных, фильтрация, запись соответствующих параметров в переменные.

Интерфейс, доступный пользователю, представляет собой веб-страницу, содержащую графическое представление характеристик микросети в зависимости от времени. Пользователю доступно меню выбора одного из регистрируемых параметров системы. Графическое

представление данных осуществляется с помощью бесплатной библиотеки Google Charts. Выбор этой библиотеки обусловлен лучшей производительностью по сравнению с аналогами, так как это важно при больших объемах обрабатываемых данных.

Страница написана на языке гипертекстовой разметки HTML, взаимодействие с данными осуществляется с помощью языка JavaScript и указанных ранее бесплатных библиотек. Для оформления используются элементы языка CSS.

3.3. Анализ работы микросети. Внедрение системы мониторинга облегчило проведение анализа работы микросети. Минимальное напряжение на блоке аккумуляторов составило 11,72 В, максимальное – 14 В. В ходе эксплуатации фиксировался заряд аккумуляторного блока. На рисунке 4 представлена ситуация, при которой в ночное время зарегистрировано снижение уровня заряда ниже допустимого и была использована внешняя сеть для зарядки аккумуляторов. На следующий день заряд восполнялся в рабочее время с помощью газогенератора.



Рис. 4. Пример использования системы мониторинга для изучения микросети

Вертикальная ось отражает заряд аккумуляторов (%), горизонтальная ось – дату и время

В ходе анализа накопленных данных было отмечено, что микросеть обеспечивает стабильную работу потребителей не только за счет использования генератора и газификатора, но и за счет отсутствия нагрузки в ночное время суток в большинстве случаев. Это существенно снижает уровень потребления. Также фиксировался стабильный уровень напряжения аккумуляторного блока (рис. 5).

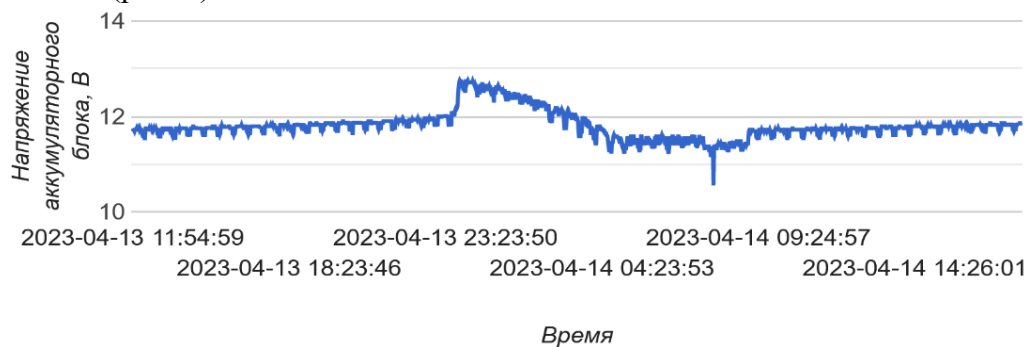


Рис. 5. Пример использования системы мониторинга для изучения микросети

Вертикальная ось отражает напряжение аккумуляторного блока, горизонтальная ось – дату и время

Ниже представлены данные из открытых источников [17] о величине облучения поверхности в точке установки панели пучком солнечных лучей (таблица 1):

Таблица 1. Данные о величине облучения панели пучком солнечных лучей

Месяц	Солнечная инсоляция, кВт*ч/м ²	Месяц	Солнечная инсоляция, кВт*ч/м ²
Январь	1,65	Июль	5,61
Февраль	2,98	Август	4,73

Март	4,43	Сентябрь	3,50
Апрель	5,29	Октябрь	2,43
Май	5,58	Ноябрь	1,85
Июнь	5,73	Декабрь	1,14

Обработка данных об инсоляции, совмещенная с данными мониторинга, позволяет вносить изменения в режимы работы установки. В месяцы низкой солнечной инсоляции требуется обеспечение работоспособности системы за счет более частого использования газификатора биомассы. Также, эти данные указывают на более высокую актуальность системы в регионах с большим количеством солнечных дней и большей величиной инсоляции относительно текущего расположения.

Основополагающей задачей при описании микросети является создание прогностической модели для различных сценариев работы. Одним из решений этой задачи является использование мониторинговых и аналитических инструментов при наличии достаточно большой базы данных.

Система мониторинга позволяет отслеживать и анализировать возникающие ситуации. Так, при мониторинге системы были замечены периоды полной или почти полной разрядки аккумуляторного блока при отсутствии солнечного света и были приняты меры по запуску электрогенераторных систем.

В ходе мониторинга были выявлены пики энергопотребления, связанные с началом и окончанием рабочего времени. Было установлено, что средняя продолжительность работы фотовольтаических элементов составляет 9 часов, согласно расположению панелей и влиянию окружающих объектов. Выявлено, что использования только солнечных элементов хватает для поддержания стабильной работы потребителей. Проведенные эксперименты показали, что использования солнечных элементов и накапливаемой энергии достаточно для запуска и поддержания работы систем газогенератора. Выявлено снижение эффективности солнечных элементов на 15% в зимний период.

Представленные аналитические данные являются основой для последующих работ, связанных с углубленным исследованием работы микросети в различных условиях и режимах.

Заключение. В ходе работы была создана система мониторинга работы микросети лабораторного масштаба. По результатам сбора и анализа данных работы системы было установлено, что при текущей конфигурации поддерживалась автономная работа и был удовлетворен спрос потребителей. За счет фотоэлектрической энергии удавалось поддерживать автономную работу в дневное время, а аккумуляторный блок накапливал избыток электроэнергии для последующей работы микросети в ночное время. При необходимости, микросеть автоматически подключалась к узлу генератора и использовала накопленную в аккумуляторах электроэнергию.

Разработанная система является примером применения современных информационных технологий для решения научных проблем.

Текущая конфигурация системы мониторинга позволяет решать задачу удаленного доступа к данным о режимах работы микросети. Однако, в последующих работах система будет дополнена новыми функциями: автоматизированным и более частым обновлением данных, улучшенной фильтрацией данных по датам, системой удаленного управления.

Благодарности. Работа выполнена при финансовой поддержке гранта Министерства науки и высшего образования Российской Федерации (проект № 075-15-2022-1215) с использованием оборудования ЦКП «Высокотемпературный контур».

Список источников

1. Hassan Q., Algburi S., Sameen A.Z., et al. A comprehensive review of international renewable energy growth. Energy and built environment, 2024, ISSN 2666-1233, DOI:10.1016/j.enbenv.2023.12.002
2. He W., Abbas Q., Alharthi M., et al. Integration of renewable hydrogen in light-duty vehicle: Nexus between energy security and low carbon emission resources. International journal of hydrogen energy, 2020, vol. 45, iss. 51, pp. 27958-27968, DOI:10.1016/j.ijhydene.2020.06.177
3. Sathish T., Agbulut Y., George S.M., et al. Waste to fuel: Synergetic effect of hybrid nanoparticle usage for the improvement of CI engine characteristics fuelled with waste fish oils. Energy, 2023, vol. 275, 127397, DOI:10.1016/j.energy.2023.127397
4. Abdin Z., Hhalaf N. Al, McGrath B. Feasibility of hydrogen hybrid energy systems for sustainable on- and off-grid integration: An Australian REZs case study. International journal of hydrogen energy, 2024, vol. 57, pp. 1197-1207, DOI:10.1016/j.ijhydene.2024.01.122
5. Харченко В.В. Микросети на основе ВИЭ: концепция, принципы построения, перспективы использования / В.В. Харченко // Энергия: экономика, техника, экология, 2014. – № 5. – С. 20-27. – EDN SEQBFP
6. Sepasi S., Toledo S., Koabayashi J., et al. A practical solution for excess energy management in a diesel-backed microgrid with high renewable penetration. Renewable energy, 2023, vol. 202, pp. 581-588, DOI:10.1016/j.renene.2022.11.114
7. Stevanato N., Pellecchia G., Sangiorgio I., et al. Planning third generation minigrids: Multi-objective optimization and brownfield investment approaches in modelling village-scale on-grid and off-grid energy systems. Renewable and sustainable energy transition, 2023, vol. 3, 100053, DOI:10.1016/j.rset.2023.100053
8. Сейфуллин А.Ю. Задачи и стратегия управления конфигурацией сельской микросети, содержащей возобновляемые источники энергии / А.Ю. Сейфуллин, А.В. Виноградов, А.В. Виноградова // Вестник Казанского государственного аграрного университета, 2021. – Т. 16, – № 3(63). – С. 90-97. – DOI:10.12737/2073-0462-2021-90-97. – EDN AHMPSE.
9. Законьшек Я. Моделирование микросети с подключением силового оборудования в реальном масштабе времени / Я. Законьшек, Ф.А. Иванов, М.А. Шамис // Энергия единой сети, 2019. – № 1(43). – С. 58-67. – EDN ZYZUJN.
10. Ajay Kumar Bansal Sizing and forecasting techniques in photovoltaic-wind based hybrid renewable energy system: A review, Journal of cleaner production, 2022, vol. 369, 133376, DOI:10.1016/j.jclepro.2022.133376
11. Afonaa-Mensah S., Odoi-Yorke F., Majeed I.B. Evaluating the impact of industrial loads on the performance of solar PV/diesel hybrid renewable energy systems for rural electrification in Ghana. Energy conversion and management: X, 2024, vol. 21, 100525, DOI:10.1016/j.ecmx.2024.100525
12. Kang H., Jung S., Kim H., et al. Multi-objective sizing and real-time scheduling of battery energy storage in energy-sharing community based on reinforcement learning. Renewable and sustainable energy reviews, 2023, vol. 185, 113655, DOI:10.1016/j.rser.2023.113655
13. Иванов Р.А. Организация мониторинга параметров экспериментального стенда солнечных панелей / Р.А. Иванов, Н. В. Максаков // Информационные и математические технологии в науке и управлении, 2021. – № 4(24). – С. 77-87. – DOI:10.38028/ESI.2021.24.4.008. – EDN IIRJJB.
14. Acresto. Informer technologies, available at: <https://acresso.software.informer.com/> (accessed: 06/09/2024)
15. Charts. Google for Developers, available at: <https://developers.google.com/chart?hl=ru> (accessed: 06/10/2024).
16. Papa Parse. Powerful CSV Parser for JavaScript, available at: <https://www.papaparse.com/> (accessed: 06/10/2024).
17. Значение солнечной инсоляции в г. Иркутск (Иркутская область). Betaenergy – URL: <https://www.betaenergy.ru/insolation/irkutsk/> (дата обращения: 10.06.2024).

Баденко Владислав Вадимович. Инженер-исследователь, аспирант отдела теплоэнергетических систем ИСЭМ СО РАН. Основные направления исследований - термохимическая конверсия биомассы, численное моделирование процессов горения и комбинированные энергетические установки для разложения биомассы. AuthorID: 1186521, SPIN: 5115-6229, ORCID: 0009-0007-0992-564X, badenko@isem.irk.ru, 664033, г. Иркутск, ул. Лермонтова 130.

Козлов Александр Николаевич. Кандидат технических наук, заведующий лабораторией термодинамики, отдел теплоэнергетических систем ИСЭМ СО РАН. Основные направления исследований - сжигание гетерогенных твердых топлив (биомасса, уголь) и разработка инженерных методов расчета процесса газификации. AuthorID: 217821, SPIN: 5451-8611, ORCID: 0000-0002-5983-5884, kozlov@isem.irk.ru, 664033, г. Иркутск, ул. Лермонтова 130.

UDC 620.91

DOI:10.25729/ESI.2024.35.3.006

Initial analysis of the performance of a laboratory-scale hybrid microgrid using a monitoring system

Vladislav V. Badenko, Alexander N. Kozlov

Melentiev Energy Systems Institute SB RAS, Russia, Irkutsk, *badenko@isem.irk.ru*

Abstract. One of the directions of the modern energy industry is the development of systems involving the utilization of solar energy, wind energy, and the recycling of secondary energy resources (waste). This paper considers an energy system, denoted as a laboratory-scale hybrid microgrid, that incorporates the use of such sources. The energy system includes solar panels, a biomass gasifier, and an electric generator. Of great importance for such systems are monitoring tools to analyze the modes of operation and make modifications. With this in mind, a monitoring system was developed to track various parameters of the laboratory equipment. Thanks to the developed system, data on the operation of the microgrid was obtained, and its limitations and weaknesses were identified.

Keywords: solar energy, biomass, monitoring

Acknowledgements: The research was funded by the Ministry of Science and Higher Education of the Russian Federation (no. 075-15-2022-1215) using the resources of the High-Temperature Circuit Multi-Access Research Center

References

1. Hassan Q., Algburi S., Sameen A.Z., et al. A comprehensive review of international renewable energy growth. Energy and built environment, 2024, ISSN 2666-1233, DOI:10.1016/j.enbenv.2023.12.002
2. He W., Abbas Q., Alharthi M., et al. Integration of renewable hydrogen in light-duty vehicle: Nexus between energy security and low carbon emission resources. International journal of hydrogen energy, 2020, vol. 45, iss. 51, pp. 27958-27968, DOI:10.1016/j.ijhydene.2020.06.177
3. Sathish T., Agbulut Y., George S.M., et al. Waste to fuel: Synergetic effect of hybrid nanoparticle usage for the improvement of CI engine characteristics fuelled with waste fish oils. Energy, 2023, vol. 275, 127397, DOI:10.1016/j.energy.2023.127397
4. Abdin Z., Hhalaf N. Al, McGrath B. Feasibility of hydrogen hybrid energy systems for sustainable on- and off-grid integration: An Australian REZs case study. International journal of hydrogen energy, 2024, vol. 57, pp. 1197-1207, DOI:10.1016/j.ijhydene.2024.01.122
5. Harchenko V.V. Mikroseti na osnove VIE: koncepciya, principy postroeniya, perspektivy ispol'zovaniya [Microgrids based on RES: concept, principles of construction, prospects of use]. Energiya: ekonomika, tekhnika, ekologiya [Energy, Economy, Technology, Ecology], 2014, no. 5, pp. 20-27.
6. Sepasi S., Toledo S., Koabayashi J., et al. A practical solution for excess energy management in a diesel-backed microgrid with high renewable penetration. Renewable energy, 2023, vol. 202, pp. 581-588, DOI: 10.1016/j.renene.2022.11.114
7. Stevanato N., Pellicchia G., Sangiorgio I., et al. Planning third generation minigrids: Multi-objective optimization and brownfield investment approaches in modelling village-scale on-grid and off-grid energy systems. Renewable and sustainable energy transition, 2023, vol. 3, 100053, DOI:10.1016/j.rset.2023.100053
8. Seyfullin A.Y., Vinogradov A.V., Vinogradova A.V. Zadachi i strategiya upravleniya konfiguraciej sel'skoj mikroseti, soderzhashchej vozobnovlyaemye istochniki energii [Objectives and strategy for managing the configuration of a rural micro-grid containing renewable energy sources]. Vestnik Kazanskogo gosudarstvennogo agrarnogo universiteta [Bulletin of Kazan state agrarian university], 2021, vol.16, no.3, pp. 90-97, DOI:10.12737/2073-0462-2021-90-97
9. Zakonshek Y., Ivanov F.A., Shamis M.A. Modelirovanie mikroseti s podklyucheniem silovogo oborudovaniya v real'nom masshtabe vremeni [Modeling of microgrid with power equipment connection in real time]. Energy of the Energiya yedinoj seti [Unified Grid], 2019, no. 1, pp. 58-67.
10. Ajay Kumar Bansal Sizing and forecasting techniques in photovoltaic-wind based hybrid renewable energy system: A review, Journal of cleaner production, 2022, vol. 369, 133376, DOI:10.1016/j.jclepro.2022.133376
11. Afonaa-Mensah S., Odoi-Yorke F., Majeed I.B. Evaluating the impact of industrial loads on the performance of solar PV/diesel hybrid renewable energy systems for rural electrification in Ghana. Energy conversion and management: X, 2024, vol. 21, 100525, DOI:10.1016/j.ecmx.2024.100525
12. Kang H., Jung S., Kim H., et al. Multi-objective sizing and real-time scheduling of battery energy storage in energy-sharing community based on reinforcement learning. Renewable and sustainable energy reviews, 2023, vol. 185, 113655, DOI:10.1016/j.rser.2023.113655

13. Ivanov R.A., Maksakov N.V. Organizaciya monitoringa parametrov eksperimental'nogo stenda solnechnyh panelj [Organization of monitoring of the parameters of the experimental stand of solar panels]. Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii [Information and mathematical technologies in science and management], 2021, no.24, pp. 77-87, DOI:10.38028/ESI.2021.24.4.008
14. Acresto. Informer technologies, available at: <https://acresso.software.informer.com/> (accessed: 06/09/2024)
15. Charts. Google for Developers, available at: <https://developers.google.com/chart?hl=ru> (accessed: 06/10/2024).
16. Papa Parse. Powerful CSV Parser for JavaScript, available at: <https://www.papaparse.com/> (accessed: 06/10/2024).
17. Znachenie solnechnoj insoljatsii v g. Irkutsk (Irkutskaja oblast') [Value of solar insolation in Irkutsk (Irkutsk region)]. Betaenergy, available at: <https://www.betaenergy.ru/insolation/irkutsk/> (accessed: 06/10/2024).

Vladislav Vadimovich Badenko. Research engineer, postgraduate student of the heat and power systems department of ISEM SB RAS. Main research areas - thermochemical conversion of biomass, numerical modelling of combustion processes and combined heat and power plants for biomass decomposition. AuthorID: 1186521, SPIN: 5115-6229, ORCID: 0009-0007-0992-564X, badenko@isem.irk.ru, 664033, Irkutsk, Lermontova St., 130.

Alexander Nikolaevich Kozlov. Senior researcher (PhD) at the department of thermal power systems at ESI SB RAS. His research interests focus on heterogeneous solid fuel (biomass, coal) combustion and the design of simple engineering techniques for the calculation of the gasification process. AuthorID: 217821, SPIN: 5451-8611, ORCID: 0000-0002-5983-5884, kozlov@isem.irk.ru, 664033, Irkutsk, Lermontova St., 130.

Статья поступила в редакцию 18.06.2024; одобрена после рецензирования 16.07.2024; принята к публикации 10.10.2024.

The article was submitted 06/18/2024; approved after reviewing 07/16/2024; accepted for publication 10/10/2024.

Элементы концепции структурно-функционального подхода к проблеме управления динамическими состояниями технических объектов в условиях вибрационных нагрузений

Елисеев Андрей Владимирович, Миронов Артем Сергеевич

Иркутский государственный университет путей сообщения,

Россия, Иркутск, *eavsh@ya.ru*

Аннотация. Развивается системный подход к проблеме оценки, формирования и коррекции динамических состояний технических объектов, находящихся в условиях вибрационных нагрузений, с помощью встроенных структурных образований. Используются методы теоретической механики, теории колебаний, теории цепей, теории автоматического управления, системного анализа. Разработаны элементы структурно-функционального подхода, заключающегося в формировании свойств структурных образований механических колебательных систем, способных выполнять определенные функции в рамках общей системы.

Ключевые слова: Механическая колебательная система, диада, структурное образование, характерные точки, частотная энергетическая функция, функция демпфирования

Цитирование: Елисеев А.В. Элементы концепции структурно-функционального подхода к проблеме управления динамическими состояниями технических объектов в условиях вибрационных нагрузений / А.В. Елисеев, А.С. Миронов // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 82-93. – DOI: 10.25729/ESI.2024.35.3.007.

Введение. В настоящее время значительный интерес вызывает развитие системных представлений о способах управления динамическими состояниями технических объектов в условиях вибрационных нагрузений с помощью структурных образований с заданными свойствами.

В рамках задач оценки, формирования и коррекции динамических состояний существенное внимание обращают на себя элементы виброзащитных систем, вибрационных технологических и транспортных машин.

Демпфирующие устройства, рассматриваемые как средства рассеивания энергии, находят своё применение в широком круге отраслей для снижения вибрации и шума, а также для контроля колебаний технических объектов, находящихся в условиях вибрационных нагрузений. Интересны возможности варьирования вязкоупругих свойств за счет начального поджатия в амортизаторах, включающих в свой состав цельнометаллические подушки [1]. Амортизатор с цельнометаллической подушкой может быть рассмотрен как структурное образование, формирующее динамические свойства системы в целом. Ключевым фактором динамических особенностей в условиях вибрационных нагрузений являются частоты колебаний элементов системы [2]. Наравне с элементами, свойства которых зависят от частоты, интерес могут представлять элементы, обладающие свойствами независимости от частот внешних возмущений [3]. Другим методом коррекции динамических состояний является использование присоединенных масс [4], к примеру, в качестве динамических гасителей.

Задачи формирования динамических состояний находят своё применение в настройке форм траекторий движений точек рабочих органов вибрационных технологических машин за счет выбора направления возмущающих сил [5]. Необходимость коррекции динамических состояний возникает при работе многокорпусных конструкций, жесткость которых определяется натягом болтовых соединений с учетом неударяющих связей [6].

Особое значение задач оценки, формирования и коррекции динамических состояний проявляется в решении проблем безопасности технических объектов железнодорожного транспорта, в частности, для оценки структурной усталости железнодорожных рельсов [7], для

снижения амплитудных пиков колебаний внутри зданий за счет использования эластичных конструктивных элементов под шпалами [8], для учета свойств композитных [9] и традиционных железобетонных шпал с учетом их взаимного расположения [10].

Проблемы безопасности и обеспечения динамического качества движения железнодорожного состава по рельсовому пути обращают внимание на возможности построения аналитических оценок динамических взаимодействий между шпалами, рельсами и колесом транспортной тележки на основе характеристик эквивалентных жесткостей [11]. В рамках тенденций роста скоростей и масс перевозимых грузов актуальна разработка методов обнаружения вышедших из строя амортизаторов в основной подвеске железнодорожного транспортного средства. В частности, в оценке динамического состояния подвески железнодорожной тележки существенное значение имеет симметрия характеристик амортизаторов и расположения точек, на основе которых производится оценка динамических состояний [12]. Важным примером характерной проблемы, связанной с техническими объектами железнодорожного транспорта в условиях интенсивных нагрузений, является нарушение контакта скользящей головки пантографа с контактным проводом [13].

Интересным примером, относящимся и к транспортным объектам, и к вибрационным технологиям, являются вибрационные транспортные средства, которые способны перемещаться в специфических средах. В частности, вибрационные нагружения на элементы механической колебательной системы способны приводить к направленному движению [14].

На основе рассмотренных вариантов технических объектов, работающих в условиях вибрационных нагружений, может быть сформулирована гипотеза, что динамическое состояние технического объекта определяется структурным образованием, реализующим определенную функцию оценки, формирования или коррекции динамического состояния системы.

Вместе с тем, вопросы разработки обобщенных моделей структурных образований в рамках структурно-функционального подхода ещё не получили широкого распространения в моделировании динамических состояний технических объектов.

Статья посвящается вопросам развития структурно-функционального подхода, ориентированного на развитие представлений о связях между структурой объекта и его функцией в механической колебательной системе.

1. Основные положения. Расчетной схемой технического объекта служит механическая колебательная система, совершающая вынужденные малые колебания под воздействием внешних возмущений кинематической или силовой природы. Обобщенные координаты системы совершают установившиеся формы колебаний в предположении, что внешние возмущения имеют моногармоническую форму. Установившиеся колебания обобщенных координат происходят на частоте внешних возмущений. Для фиксированной частоты внешних возмущений движения системы представляются конечной совокупностью амплитуд колебаний обобщенных координат.

В качестве структурного образования рассматривается диада – система с двумя степенями свободы, включающая в свой состав два массоинерционных элемента, соединенных упругим элементом с учетом возможных дополнительных связей (рис. 1), встраиваемых параллельно упругому элементу [15-17].

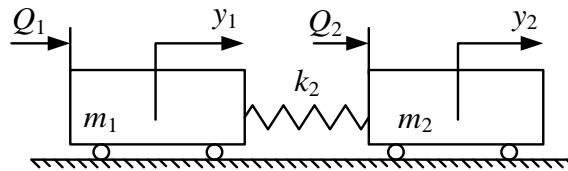


Рис. 1. Диада. m_1, m_2 – массоинерционные элементы, k_2 – упругий элемент, Q_1, Q_2 – внешние силовые возмущения, y_1, y_2 – обобщенные координаты

Динамическое, зависящее от частоты внешнего возмущения, состояние диады определяется двумя амплитудами колебаний обобщенных координат. В общем случае возможны различные варианты диады в зависимости от дополнительных связей, особенности которых определяются учетом сил трения и характеристиками массоинерционных коэффициентов устройств для преобразования движений.

Диада, включенная в механическую колебательную систему, обладает потенциалом оценки, формирования и коррекции динамических состояний включающей системы. Выбор параметров диады может быть реализован таким образом, чтобы диада, рассматриваемая как структурное образование в рамках системы, выполняла определенную функцию.

Задача исследования заключается в разработке методологических основ структурно-функционального подхода, отличающегося тем, что структурному образованию сопоставляется определенная функция в рамках рассматриваемой системы.

2. Структурная математическая модель. На основе известных методов расчетной схеме (рис. 1) сопоставляется структурная схема (рис. 2) эквивалентной в динамическом отношении системы автоматического управления [18].

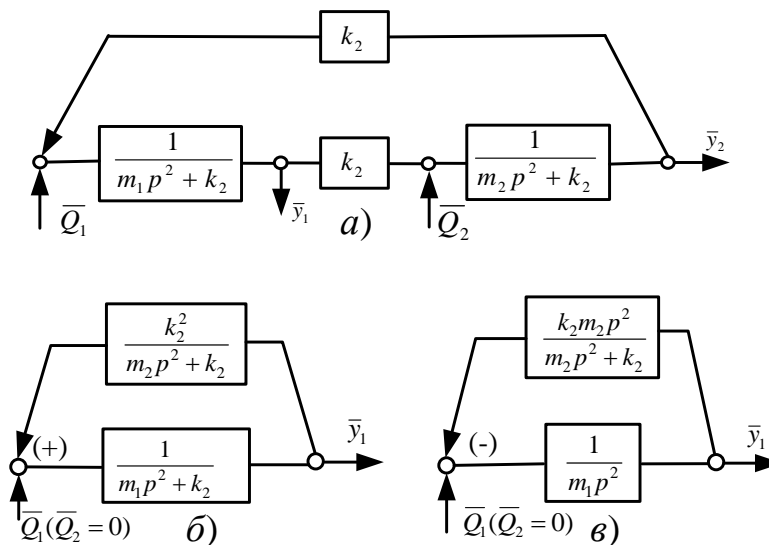


Рис. 2. Структурное образование в виде диады: *а)* структурная схема; *б)* структурная схема с исключенной координатой \bar{y}_2 при $\bar{Q}_2 = 0$; *в)* – структурная схема с выделением объектом m_1 , охваченным контуром обратной связи;

Динамические состояния диады могут быть охарактеризованы на основе амплитудно-частотных характеристик передаточных функций системы и передаточных функций межпарциальных связей [19].

Амплитудно-частотные характеристики передаточных функций отображают основные динамические особенности системы в виде частот динамического гашения, обнуляющих амплитуды установившихся колебаний обобщенных координат системы, резонансы и интервалы частот, обеспечивающих синфазные и противофазные формы взаимодействий между парци-

альными системами и внешними возмущениями (рис. 3). Объединение одиночных воздействий на диаду приводит к формированию сложных динамических состояний в виде режимов динамического гашения, частоты которых зависят от конфигурации внешних воздействий [20].

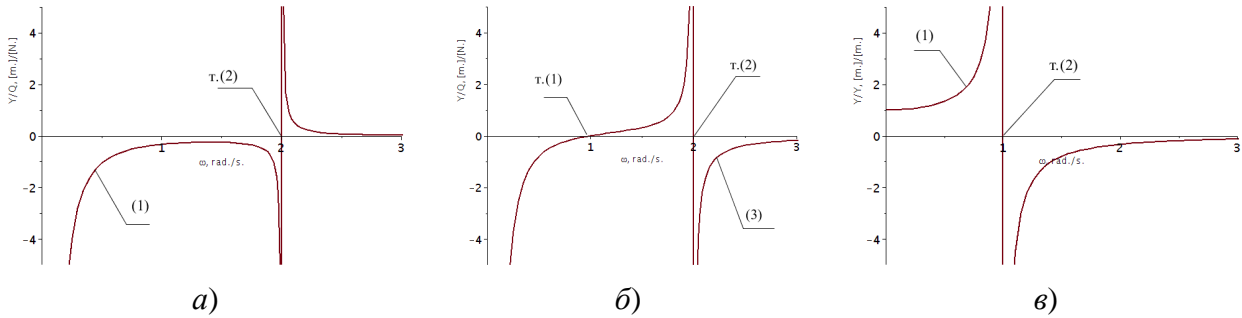


Рис.3. Особенности динамических состояний диады: а) амплитудно-частотная характеристики передаточной функции диады без режима динамического гашения колебаний; б) амплитудно-частотная характеристика передаточной функции диады с учетом режима динамического гашения колебаний; в) амплитудно-частотная характеристика передаточной функции межпарциальных связей диады

3. Определение характерных точек диады с учетом устройства для преобразования движения. В общем случае диада может включать дополнительные упругие связи, амортизаторы и устройства для преобразования движения, в совокупности рассматриваемые как устройства регулирования динамических состояний [21].

Для диады, совершающей свободные или вынужденные колебания, могут быть определены характерные точки, которые удовлетворяют определенным условиям, в частности, отображают свойства неподвижности в определенной системе координат, или свойства динамической симметрии; характерные точки могут совпадать с центром масс [22].

Способом изменения характерных точек диады является включение звена с передаточной функцией второго дифференцирования Lp^2 параллельно упругому элементу (рис. 4). Варьирование параметров способно изменять границы пространственной реализации диады.

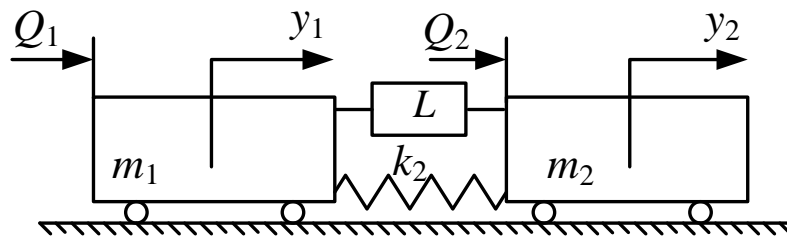


Рис. 4. Диада, включающая в состав устройство для преобразования движения L

Изменение массоинерционного коэффициента L устройства для преобразования движения определяет положение неподвижной точки относительно предельных положений массоинерционных элементов диады (рис. 5а), представленных обобщенными координатами с учетом длины упругого элемента [22].

Положение характерной точки диады обладает чувствительностью к изменению коэффициента L , принимающей отрицательные, нулевые и положительные значения в зависимости от массоинерционных элементов диады (рис. 5б).

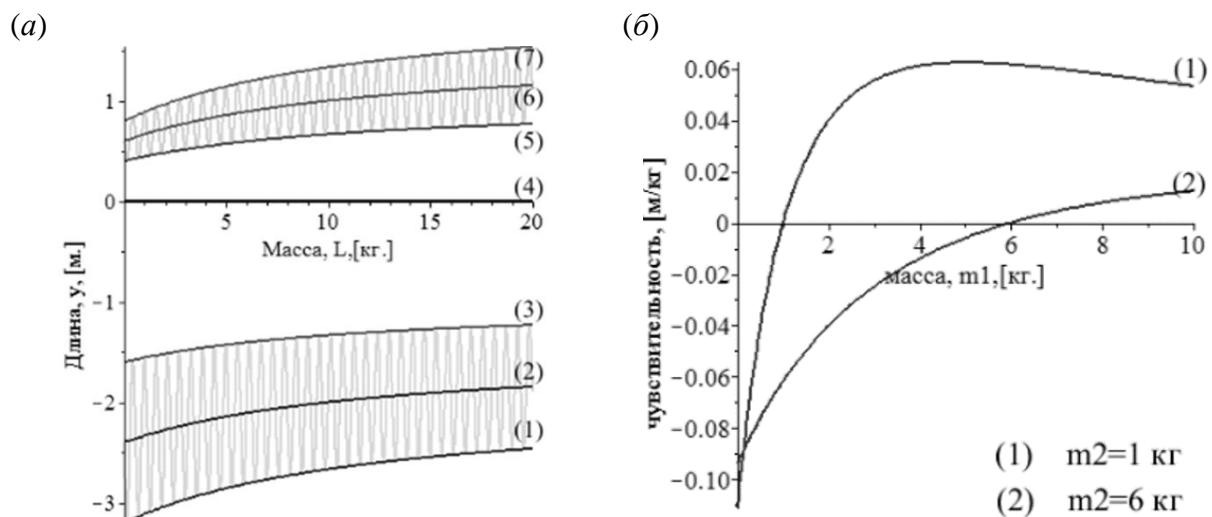


Рис.5. Особенности динамических состояний диады в зависимости от массоинерционного параметра L устройства преобразования движения ($m_1 < m_2$): а) приведение семейства диад к фиксированному положению неподвижной характерной точки; б) график чувствительности положения характерной точки по отношению к параметру L на интервале $m_1 \in [0..10]$, линия (1) $m_2=1$ кг., линия (2) $m_2=6$ кг.

4. Оценка динамических состояний диады на основе энергетических соотношений.

Наравне с оценкой динамических состояний с помощью амплитудно-частотных характеристик диады могут быть охарактеризованы с помощью частотных энергетических функций и функций демпфирования [23]. В зависимости от варьирования параметра системы, к примеру, массоинерционного коэффициента устройства преобразования движения L или коэффициента отношения масс диады, с помощью параметризующей функции может быть построена совокупность частотных энергетических функций и функций демпфирования, существенно изменяющихся для различных параметров системы (рис. 6).

Полная совокупность динамических состояний диад в свободных формах движения может быть охарактеризована частотными энергетическими функциями и функциями демпфирования, отображающими особенности учета сил вязкого трения в системе [24]. Частотные энергетические функции, являясь аналогом отношения Рэлея [25], для систем без учета трения отображают условие равенства максимальных значений кинетической и потенциальной энергии, которое позволяет выразить частоту в зависимости от формы колебания элементов механической колебательной системы. Частотная энергетическая функция, рассматриваемая как функция формы колебания системы, в экстремальных точках достигает значений собственных частот, а в критических точках (ноль или бесконечность) – значений парциальных частот.

Особенности частотных энергетических функций позволяют рассматривать собственные колебания, как колебания системы, для которых максимальные значения потенциальной и кинетической энергии равны между собой. В свою очередь, вынужденные колебания системы могут рассматриваться как колебания, для которых максимальные значения потенциальной и кинетической энергии различны. Силы трения могут быть учтены путем рассмотрения модифицированной частотной энергетической функции и функции демпфирования, которые также в экстремальных значениях достигают значений собственных частот и коэффициентов затухания колебаний соответственно.

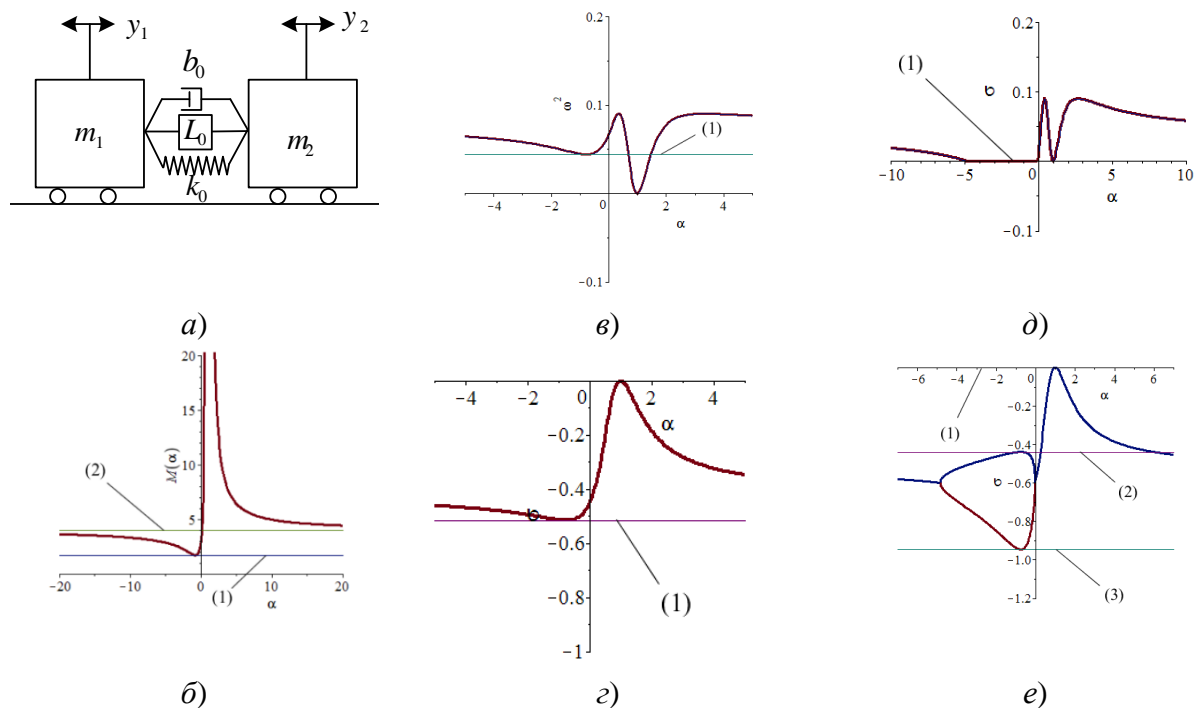


Рис. 6. Функции для анализа динамических особенностей диады: *a* – расчетная схема диады; *б* – параметризирующая функция; *в* – частотная энергетическая функция (вариант 1); *г* – функция демпфирования (вариант 1); *д* – частотная энергетическая функция (вариант 2); *е* – функция демпфирования (вариант 2)

Для анализа энергетических особенностей механическая колебательная система может быть преобразована к системе уравнений относительно рычажных и упругих связей.

5. Отображение динамических особенностей диады с помощью упругих и рычажных связей. На основе расчетной схемы диады (рис. 1) с помощью формализма Лагранжа и использования интегрального преобразования Лапласа с учетом нулевых начальных условий может быть построена система линейных уравнений относительно изображений для вариантов кинематических (Табл1, строка 1, столбец I) и силовых возмущений (Табл1, строка 1, столбец II).

Для учета форм колебаний координат системы и связности внешних возмущений могут быть введены коэффициенты $\alpha_1, \beta_1, \gamma_1, \gamma_2$ (табл.1, строка 2) отображающие рычажные связи между обобщенными координатами и силами системы, а также коэффициент β_2 (табл.1, строка 2), отображающий динамическую податливость.

В рамках новых обозначений линейные системы в изображениях (табл. 1, строка 1) с учетом системы связей (табл. 1, строка 2) приводятся к системам относительно коэффициентов рычажных связей и коэффициента податливости (табл. 1, строка 3).

Для отображение характерных особенностей рычажных связей, расчетная схема (рис. 1) может быть представлена в эквивалентном в виде с помощью виртуальных рычагов (рис. 7) [26].

В частности, диада с учетом наличия связи с опорными поверхностям в условиях кинематического возмущения может быть представлена в эквивалентном виде (рис. 7а), где связанные кинематические возмущения представлены виртуальным рычагом A_1B_1 . Связные кинематические возмущения приводят к движению массоинерционных элементов диады с амплитудами, которые также могут быть представлены с помощью виртуального рычага AB , обладающим характерной точкой, являющейся неподвижной точкой опоры $t.O$.

Таблица 1. Аналитические характеристики диады в условиях кинематических и силовых возмущений

	I	II
1	$\begin{pmatrix} m_1 p^2 + k_0 + k & -k_0 \\ -k_0 & m_2 p^2 + k_0 + k \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} = \begin{pmatrix} k \bar{z}_1 \\ k \bar{z}_2 \end{pmatrix}$	$\begin{pmatrix} m_1 p^2 + k_0 & -k_0 \\ -k_0 & m_2 p^2 + k_0 \end{pmatrix} \begin{pmatrix} \bar{y}_{01} \\ \bar{y}_{02} \end{pmatrix} = \begin{pmatrix} \bar{Q}_1 \\ \bar{Q}_2 \end{pmatrix}$
2	$z_2 = \gamma_1 z_1, \bar{y}_2 = \alpha_1 \bar{y}_1, \bar{y}_1 = \beta_1 \bar{z}_1$	$Q_2 = \gamma_2 Q_1, \bar{y}_{01} = \beta_2 \bar{Q}_1, \bar{y}_{02} = \alpha_2 \bar{y}_{01}$
3	$\begin{pmatrix} m_1 p^2 + k_0 + k & -k_0 \\ -k_0 & m_2 p^2 + k_0 + k \end{pmatrix} \begin{pmatrix} \beta_1 \\ \alpha_1 \beta_1 \end{pmatrix} = \begin{pmatrix} k \\ k \gamma_1 \end{pmatrix}$	$\begin{pmatrix} -m_1 \omega^2 + k_0 & -k_0 \\ -k_0 & -m_2 \omega^2 + k_0 \end{pmatrix} \begin{pmatrix} \beta_2 \\ \alpha_2 \beta_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \gamma_2 \end{pmatrix}$
4	$\alpha_1 = \frac{\begin{vmatrix} -m_1 \omega^2 + k_0 + k & k \\ -k_0 & k \gamma_1 \end{vmatrix}}{\begin{vmatrix} k & -k_0 \\ k \gamma_1 & -m_2 \omega^2 + k_0 + k \end{vmatrix}}$	$\alpha_2 = \frac{\begin{vmatrix} -m_1 \omega^2 + k_0 & 1 \\ -k_0 & \gamma_2 \end{vmatrix}}{\begin{vmatrix} 1 & -k_0 \\ \gamma_2 & -m_2 \omega^2 + k_0 \end{vmatrix}}$
	$\beta_1 = \frac{\begin{vmatrix} k & -k_0 \\ k \gamma_1 & -m_2 \omega^2 + k_0 + k \end{vmatrix}}{\begin{vmatrix} -m_1 \omega^2 + k_0 + k & -k_0 \\ -k_0 & -m_2 \omega^2 + k_0 + k \end{vmatrix}}$	$\beta_2 = \frac{\begin{vmatrix} 1 & -k_0 \\ \gamma_2 & -m_2 \omega^2 + k_0 \end{vmatrix}}{\begin{vmatrix} -m_1 \omega^2 + k_0 & -k_0 \\ -k_0 & -m_2 \omega^2 + k_0 \end{vmatrix}}$

Вынужденные движение диады в условиях силовых возмущений аналогично могут быть представлены виртуальным рычагом A_2B_2 с опорой в т. O_2 (рис. 7б).

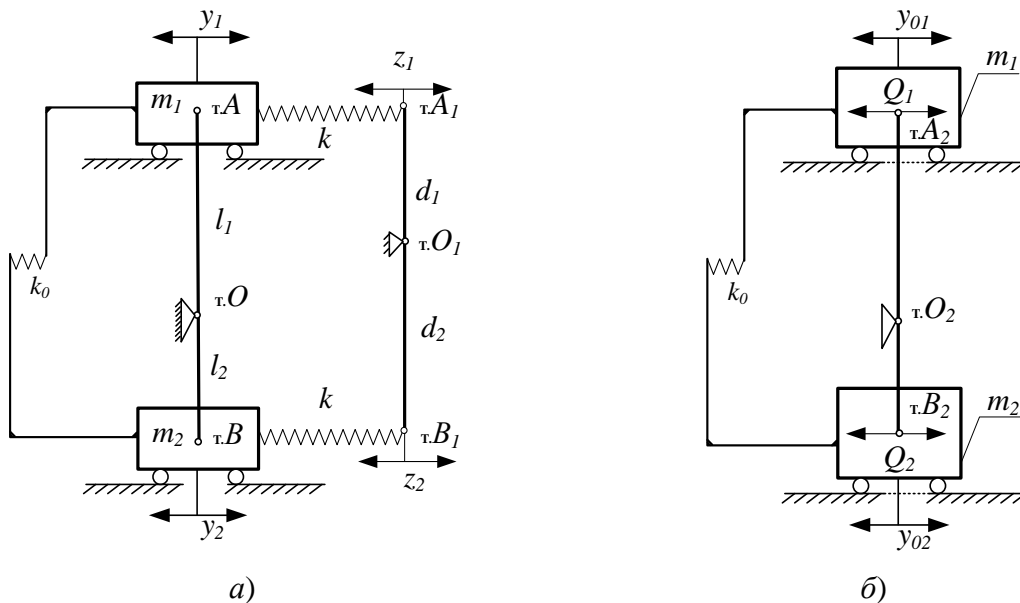


Рис.7. Отображение динамических свойств диад с помощью рычажных и упругих связей.

а – в условиях кинематических нагружений; *б* – в условиях силовых нагружений

Представления динамических состояний диад с помощью рычажных связей создает предпосылки для детализированного анализа особенностей энергетических отношений механической колебательной системы.

Разработанная система представлений может быть использована для создания способов и устройств регулирования характеристик вибрационных полей технологических машин на основе учета распределения амплитуд [26-28].

Заключение. Структурные образования имеют существенное значение в формировании динамических состояний сложных технических объектов. В системе технического объекта структурное образование способно выполнять определенные функции: регулировать форму

движения рабочего органа; сохранять неударживающий контакт, регулировать жесткости многокорпусных конструкций, изменять скорости вибротранспортных систем и др.

В качестве структурного образования, потенциально реализующего совокупность функции, выступает диада – механическая колебательная система, образованная двумя массоинерционными элементами с учетом дополнительных вязко-упругих элементов и устройств для преобразования движения.

Диада с учетом включения в состав вязко-упругих элементов и устройств для преобразования движения обладает набором настроечных параметров. Одним из настроечных параметров диады может служить массоинерционный коэффициент устройства для преобразования движения. Особенности диад могут быть отображены с помощью положений характерных точек, удовлетворяющих системе условий. Варьирование настроечных параметров способно изменять положения характерных точек и пространственных габаритов диады.

Система уравнений относительно обобщенных координат и сил может быть преобразована к системе уравнений относительно рычажных и упругих связей. На основе преобразованной системы уравнений относительно рычажных и упругих связей могут быть получены энергетические соотношения, определяющие частотную энергетическую функцию и функцию демпфирования. Множество структурных образований в виде диады может быть представлено совокупностью пар частотной энергетической функции и функции демпфирования.

Множество образований в виде диады с различными параметрами может быть представлено совокупностью пар частотных энергетических функций и функций демпфирования, представляющих собой функциональные характеристики системы.

Совокупность пар частотной энергетической функции и функции демпфирования обладает рядом топологических особенностей, которые могут быть использованы в качестве признаков классификации динамических особенностей диад.

Можно полагать, что совокупности с частотными энергетическими функциями и функциями демпфирования позволяют производить анализ энергетических соотношений свободных и вынужденных колебаний на новом уровне детализации.

Совокупность разработанных положений может быть рассмотрена как основа научно-методологической концепции структурно-функционального подхода, в рамках которого диада выступает в роли структурного образования, обладающего потенциалом реализации широкого набора функций в рамках системы за счет варьирования набора настроечных параметров, а также с помощью частоты и структуры внешних возмущений силовой или кинематической природы.

Список источников

1. Riess Sebastian, Kaal William, Herath Kristian. Frequency-Adaptable tuned mass damper using metal cushions. *Vibration*, 2021, no. 4, pp. 77-90, DOI:10.3390/vibration4010007
2. Vellingiri Vishwa, Sadasivam Udhayakumar Effect of vibrator parameters and physical characteristics of parts on conveying velocity. *Strojniški vestnik - Journal of mechanical engineering*, 2023, no 69, pp. 352-363, DOI:10.5545/sv-jme.2022.510.
3. Kletschkowski T. Theoretical and non-dimensional investigations into vibration control using viscoelastic and endochronic elements. *Vibration*, 2023, no. 6(4):1030-1047, DOI:10.3390/vibration6040060
4. Tophoj Laust, Grathwol Nikolaj, Hansen Svend. Effective mass of tuned mass dampers. *Vibration*, 2018, no 1, pp. 192-206, DOI:10.3390/vibration1010014.
5. Korendiy V., Kachur O., Predko R. et al. Generating rectilinear, elliptical, and circular oscillations of a single-mass vibratory system equipped with an enhanced twin crank-type exciter. *Vibroengineering Procedia*, 2023, no. 51, pp. 8-14, DOI:10.21595/vp.2023.23657.
6. Krot Pavlo, Hamid Shiri, Przemysław Dąbek, Radosław Zimroz Diagnostics of bolted joints in vibrating screens based on a multi-body dynamical model. *Materials*, 2023, 16, no. 17: 5794, DOI:10.3390/ma16175794.

7. Moueddeb Maryam, Louf François, Boucard Pierre-Alain, et al. An efficient numerical model to predict the mechanical response of a railway track in the low-frequency range. *Vibration*, 2022, no 5, pp 326-343, DOI:10.3390/vibration5020019.
8. Sehner Michael, Nava Luis, Seidl-Nigsch Markus, Loy Harald. Vibration mitigation: under-ballast mats in heavy-haul applications. *Practice periodical on structural design and construction*, 2023, no. 28, 05023004, DOI:10.1061/PPSCFX.SCENG-1258.
9. Zhao Zhenhang, Gao Ying, Li Chenghui Research on the vibration characteristics of a track's structure considering the viscoelastic properties of recycled composite sleepers. *Applied sciences*, 2020, no. 11, 150 p., DOI:10.3390/app11010150.
10. Zou Yu, Wen Yongpeng, Sun Qian Study on the urban rail transit sleeper spacing considering vehicle system. *MATEC Web of Conferences*, 2019, 296, 01008, DOI:10.1051/mateconf/201929601008.
11. Yoshino Shun, Abe Kazuhisa, Koro Kazuhiro An analytic solution of mathematical expectation for bogie-track interaction problems. *Mechanical engineering journal*, 2023, no. 10, DOI: 10.1299/mej.22-00300.
12. Dumitriu Madalina, Apostol Ioana Influence of interference between vertical and roll vibrations on the dynamic behaviour of the railway bogie. *Vibration*, 2022, no. 5, pp. 659-675, DOI:10.3390/vibration5040039.
13. Chu Song Yang Study on dynamic interaction of railway pantograph–catenary including reattachment momentum impact. *Vibration*, 2020, no. 3, pp. 18-33, DOI:10.3390/vibration3010003.
14. Korendiy Vitaliy, Volodymyr Gursky, Oleksandr Kachur, Petro Dmyterko, Oleh Kotsiumbas, Oleksandr Havrylchenko Mathematical model and motion analysis of a wheeled vibro-impact locomotion system. *Vibroengineering PROCEDIA*, April 21, 2022, 77–83, DOI:10.21595/vp.2022.22422.
15. Елисеев А.В. Диады в механических системах: особенности динамических свойств. Часть I / А.В. Елисеев // Вестник Иркутского государственного технического университета, 2017. – Т. 21. – № 7(126). – С. 26–38. – DOI:10.21285/1814-3520-2017-7-26-38
16. Елисеев А.В. Диады в механических системах: особенности динамических свойств (Часть II) / Елисеев А.В. // Вестник Иркутского государственного технического университета, 2017. – Т. 21. – № 8. – С. 22–37.
17. Елисеев С.В. Устройства для преобразования движения в структуре диады механической колебательной системы / С.В. Елисеев, А.И. Орленко, Д.Х. Нгуен // Вестник Донского государственного технического университета, 2017. – Т. 17. – № 3 (90). – С. 46-59.
18. Eliseev S.V., Eliseev A.V. Theory of oscillations. structural mathematical modeling in problems of dynamics of technical objects. Series: studies in systems, decision and control. Springer international publishing, Cham, 2020, vol.252, 521 p.
19. Eliseev A.V. Structural mathematical modeling applications in technological machines and transportation vehicles. Hershey, PA: IGI Global, 2023, DOI:10.4018/978-1-6684-7237-8
20. Хоменко А.П. Структурные образования в механических колебательных системах: диада как форма взаимодействия элементов / А.П. Хоменко, С.В. Елисеев // Современные технологии. Системный анализ. Моделирование, 2017. – № 2 (54). – С. 8-14.
21. Елисеев С.В. Структурные образования в механических колебательных системах: диады, их свойства, возможности изменения динамических состояний / С.В. Елисеев, А.И. Орленко, А.В. Елисеев // Транспорт Урала, 2017. – № 3 (54). – С. 56-63. – DOI:10.20291/1815-9400-2017-3-56-63
22. Елисеев А.В. Особенности динамических свойств диады в подходах с позиций мезомеханики / А.В. Елисеев, С.В. Елисеев // Современные технологии. Системный анализ. Моделирование, 2019. – № 2 (62). – С. 23-33.
23. Елисеев А.В. Частотная функция и функция демпфирования в оценке динамических процессов в механических колебательных системах с симметрией / А.В. Елисеев // *Advanced engineering research (Rostov-on-Don)*, 2020. – Т. 20. – № 4. – С. 360-369.
24. Елисеев А.В. Связность движений в системах с диссипацией энергии: системные подходы / Елисеев А.В. // Информационные и математические технологии в науке и управлении, 2020. – № 3 (19). – С. 43-56.
25. Рэлей Д.В. Теория звука. Т. 1. / Джон Вильям Стрэтт Рэлей. – Москва; Ленинград : Гостехтеоретиздат, 1940. – 500 с.
26. Елисеев А.В. Некоторые возможности управления одномерным вибрационным полем технологической машины / А.В. Елисеев, К.Ч. Вьонг // Современные технологии. Системный анализ. Моделирование, 2016. – № 1 (49). – С. 33-41.
27. Елисеев А.В. некоторые подходы к обоснованию схемы инерционного возбуждения в технологических вибрационных машинах / А.В. Елисеев, Л.А. Мамаев, И.С. Ситов // Системы. Методы. Технологии. 2015. – № 4 (28). – С. 15-24.

28. Елисеев А.В. Характерные особенности распределения амплитуд в свободных колебаниях механических систем: введение дополнительных связей, структурные образования / А.В. Елисеев, Н.К. Кузнецов, А.В. Николаев // Машиностроение: инновационные аспекты развития: Материалы международной научно-практической конференции. – Санкт-Петербург: СПбФ НИЦ МС, 2019. – №2. – 237 с.

Елисеев Андрей Владимирович. Кандидат технических наук, доцент кафедры математики, Иркутский государственный университет путей сообщения, Системный анализ механических колебательных систем. Структурное математическое моделирование. Вибрационные взаимодействия элементов технологических систем. Неудерживающие связи. AuthorID: 693903, SPIN: 8781-1123, ORCID: 0000-0003-0222-2507, eavsh@ya.ru, 664074. г. Иркутск, ул. Чернышевского 15.

Миронов Артем Сергеевич. Соискатель, Иркутский государственный университет путей сообщения. Системный анализ механических колебательных систем. AuthorID: 703296, SPIN: 6733-5363, ORCID: 0000-0002-0921-0915, art.s.mironov@mail.ru, 664074. г. Иркутск, ул. Чернышевского 15.

UDC 519.71, 629.4.015, 62-752, 534.015

DOI:10.25729/ESI.2024.35.3.007

Elements of the concept of a structural and functional approach to the problem of controlling the dynamic states of technical objects under vibration loads

Andrey V. Eliseev, Artyom S. Mironov

Irkutsk state transport university, Russia, Irkutsk, eavsh@ya.ru

Abstract. A systematic approach is being developed to the problem of assessing, forming and correcting the dynamic states of technical objects under vibration loads using embedded structural formations. The methods of theoretical mechanics, oscillation theory, circuit theory, automatic control theory, and system analysis are used. Elements of a structural and functional approach have been developed, which consists in the formation of properties of structural formations of mechanical oscillatory systems capable of performing certain functions within the framework of a common system.

Keywords: Mechanical oscillatory system, dyad, structural formation, characteristic points, frequency energy function, damping function

References

1. Riess Sebastian, Kaal William, Herath Kristian. Frequency-Adaptable tuned mass damper using metal cushions. *Vibration*, 2021, no. 4, pp. 77-90, DOI:10.3390/vibration4010007
2. Vellingiri Vishwa, Sadasivam Udhayakumar Effect of vibrator parameters and physical characteristics of parts on conveying velocity. *Strojniški vestnik - Journal of mechanical engineering*, 2023, no 69, pp. 352-363, DOI:10.5545/sv-jme.2022.510.
3. Kletschkowski T. Theoretical and non-dimensional investigations into vibration control using viscoelastic and endochronic elements. *Vibration*, 2023, no. 6(4):1030-1047, DOI:10.3390/vibration6040060
4. Tophøj Laust, Grathwol Nikolaj, Hansen Svend. Effective mass of tuned mass dampers. *Vibration*, 2018, no 1, pp. 192-206, DOI:10.3390/vibration1010014.
5. Korendiy V., Kachur O., Predko R. et al. Generating rectilinear, elliptical, and circular oscillations of a single-mass vibratory system equipped with an enhanced twin crank-type exciter. *Vibroengineering Procedia*, 2023, no. 51, pp. 8-14, DOI:10.21595/vp.2023.23657.
6. Krot Pavlo, Hamid Shiri, Przemysław Dąbek, Radosław Zimroz Diagnostics of bolted joints in vibrating screens based on a multi-body dynamical model. *Materials*, 2023, 16, no. 17: 5794, DOI:10.3390/ma16175794.
7. Moueddeb Maryam, Louf François, Boucard Pierre-Alain, et al. An efficient numerical model to predict the mechanical response of a railway track in the low-frequency range. *Vibration*, 2022, no 5, pp 326-343, DOI:10.3390/vibration5020019.

8. Sehner Michael, Nava Luis, Seidl-Nigsch Markus, Loy Harald. Vibration mitigation: under-ballast mats in heavy-haul applications. Practice periodical on structural design and construction, 2023, no. 28, 05023004, DOI:10.1061/PPSCFX.SCENG-1258.
9. Zhao Zhenhang, Gao Ying, Li Chenghui Research on the vibration characteristics of a track's structure considering the viscoelastic properties of recycled composite sleepers. Applied sciences, 2020, no. 11, 150 p., DOI:10.3390/app11010150.
10. Zou Yu, Wen Yongpeng, Sun Qian Study on the urban rail transit sleeper spacing considering vehicle system. MATEC Web of Conferences, 2019, 296, 01008, DOI:10.1051/mateconf/201929601008.
11. Yoshino Shun, Abe Kazuhisa, Koro Kazuhiro An analytic solution of mathematical expectation for bogie-track interaction problems. Mechanical engineering journal, 2023, no. 10, DOI: 10.1299/mej.22-00300.
12. Dumitriu Madalina, Apostol Ioana Influence of interference between vertical and roll vibrations on the dynamic behaviour of the railway bogie. Vibration, 2022, no. 5, pp. 659-675, DOI:10.3390/vibration5040039.
13. Chu Song Yang Study on dynamic interaction of railway pantograph–catenary including reattachment momentum impact. Vibration, 2020, no. 3, pp. 18-33, DOI:10.3390/vibration3010003.
14. Korendiy Vitaliy, Volodymyr Gursky, Oleksandr Kachur, Petro Dmyterko, Oleh Kotsiumbas, Oleksandr Havrylchenko Mathematical model and motion analysis of a wheeled vibro-impact locomotion system. Vibroengineering PROCEDIA, April 21, 2022, 77–83, DOI:10.21595/vp.2022.22422.
15. Eliseev A.V. Diady v mekhanicheskikh sistemakh: osobennosti dinamicheskikh svoistv. Chast' I [Dyads in mechanical systems: features of dynamic properties. Part I]. Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta [Bulletin of Irkutsk state technical university], 2017, vol. 21, no. 7(126), pp. 26-38.
16. Eliseev A.V. Diady v mekhanicheskikh sistemakh: osobennosti dinamicheskikh svoistv. Chast' II [Dyads in mechanical systems: features of dynamic properties. Part II]. Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta [Bulletin of Irkutsk state technical university], 2017, vol. 21, no. 8, pp. 22-37.
17. Eliseev S.V., Orlenko A.I., Nguyen D.H. Ustroistva dlya preobrazovaniya dvizheniya v strukture diady mekhanicheskoi kolebatel'oi sistemy [Devices for converting motion in the dyad structure of a mechanical oscillatory system]. Vestnik Donskogo gosudarstvennogo tekhnicheskogo universiteta [Bulletin of the Don state technical university], 2017, Level 17, no. 3 (90), pp. 46-59.
18. Eliseev S.V., Eliseev A.V. Theory of oscillations. structural mathematical modeling in problems of dynamics of technical objects. Series: studies in systems, decision and control. Springer international publishing, Cham, 2020, vol.252, 521 p.
19. Eliseev A.V. Structural mathematical modeling applications in technological machines and transportation vehicles. Hershey, PA: IGI Global, 2023, DOI:10.4018/978-1-6684-7237-8
20. Khomenko A.P., Eliseev S.V. Strukturnye obrazovaniya v mekhanicheskikh kolebatel'nykh sistemakh: diada kak forma vzaimodeistviya elementov [Structural formations in mechanical oscillatory systems: dyad as a form of interaction of elements]. Sovremennyye tekhnologii. Sistemnyy analiz. Modelirovaniye [Modern technologies. System analysis. Modeling], 2017, no. 2 (54), pp. 8-14.
21. Eliseev S.V., Orlenko A.I., Eliseev A.V. Strukturnye obrazovaniya v mekhanicheskikh kolebatel'nykh sistemakh: diady, ikh svoistva, vozmozhnost izmeneniya dinamicheskikh sostoyanii [Structural formations in mechanical oscillatory systems: dyads, their properties, the possibility of changing dynamic states]. Transport Urala [Transport of the Urals], 2017, no. 3 (54), pp. 56-63.
22. Eliseev A.V., Eliseev S.V. Osobennosti dinamicheskikh svoistv diady v podkhodakh s pozitsii mezomekhaniki [Features of dynamic properties of the dyad in approaches from the standpoint of mesomechanics]. Sovremennyye tekhnologii. Sistemnyy analiz. Modelirovaniye [Modern technologies. System analysis. Modeling], 2019, no. 2 (62), pp. 23-33.
23. Eliseev A.V. Chastotnaya funktsiya i funktsiya dempfirovaniya v otsenke dinamicheskikh protsessov v mekhanicheskikh kolebatel'nykh sistemakh s simmetriei [Frequency function and damping function in the assessment of dynamic processes in mechanical oscillatory systems with symmetry]. Advanced engineering research (Rostov-on-Don), 2020, vol. 20, no. 4, pp. 360-369.
24. Eliseev A.V. Svyaznost' dvizhenii v sistemakh s dissipatsiei energii: sistemnye podkhody [Connectivity of movements in systems with energy dissipation: systemic approaches]. Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii [Information and mathematical technologies in science and management], 2020, no. 3(19), pp. 43-56.
25. Rayleigh D.V. Teoriya zvuka. T. 1 [Theory of Sound. Vol. 1]. Moscow, Leningrad, Gostekhteorizdat, 1940, 500 p.

26. Eliseev A.V., Vyong K.C. Nekotorye vozmozhnost upravleniya odnomernym vibratsionnym polem tekhnologicheskoi mashiny [Some possibilities of controlling the one-dimensional vibrational field of a technological machine]. *Sovremennyye tekhnologii. Sistemnyy analiz. Modelirovaniye* [Modern technologies. System analysis. Modeling], 2016, no. 1(49), pp. 33-41.
27. Eliseev A.V., Mamaev L.A., Sitov I.S. Nekotorye podkhody k obosnovaniyu skhemy inertsiionnogo vzbuzhdeniya v tekhnologicheskikh vibratsionnykh mashinakh [Some approaches to substantiating the scheme of inertial excitation in technological vibrating machines]. *Sistemy. Metody. Tekhnologii* [The system. Methods. Technologies], 2015, no. 4 (28), pp. 15-24.
28. Eliseev A.V., Kuznetsov N.K., Nikolaev A.V. Kharakternye osobennosti raspredeleniya amplitud v svobodnykh kolebaniyakh mekhanicheskikh sistem: vvedenie dopolnitel'nykh svyazei, strukturnye obrazovaniya [Characteristic features of amplitude distribution in free vibrations of mechanical systems: introduction of additional bonds, structural formations]. *Mashinostroyeniye: innovatsionnyye aspekty razvitiya: Materialy mezhdunarodnoy nauchno-prakticheskoy konferentsii*. – Sankt-Peterburg: SPbF NITS MS [Mechanical engineering: innovative aspects of development: Proceedings of the International Scientific and Practical conference, St. Petersburg, SPbF SIC MS], 2019, no. 2, 237 p.

Eliseev Andrey Vladimirovich. *Candidate of technical sciences associate professor of the department of mathematics, Irkutsk state transport university. System analysis of mechanical oscillatory systems. Structural mathematical modeling. Vibrational interactions of elements of technological systems. Unstoppable connections. AuthorID: 693903, SPIN: 8781-1123, ORCID: 0000-0003-0222-2507, eavsh@ya.ru 664074. Irkutsk, Chernyshevsky str. 15.*

Artyom Sergeevich Mironov. *Applicant, Irkutsk state transport university. System analysis of mechanical oscillatory systems. AuthorID: 703296, SPIN: 6733-5363, ORCID: 0000-0002-0921-0915, art.s.mironov@mail.ru 664074. Irkutsk, Chernyshevsky str. 15.*

Статья поступила в редакцию 14.03.2024; одобрена после рецензирования 02.09.2024; принята к публикации 08.10.2024.

The article was submitted 03/14/2024; approved after reviewing 09/02/2024; accepted for publication 10/08/2024.

Моделирование трехтактного частотно-импульсного преобразователя

Дунаев Михаил Павлович, Саушкин Алексей Анатольевич

Иркутский национальный исследовательский технический университет,
Россия, Иркутск, *mdunaev10@mail.ru*

Аннотация. В статье представлена функциональная схема трехтактного частотно-импульсного преобразователя (ЧИП). Разработаны имитационные модели трехтактного ЧИП и системы электропривода ЧИП-Д в среде Matlab/Simulink. Проведено моделирование трехтактного ЧИП и системы электропривода ЧИП-Д. Приведено сравнение результатов моделирования и расчета электромеханических характеристик системы электропривода ЧИП-Д. Показано, что построенная авторами математическая модель ЧИП-Д с достаточной точностью описывает статические режимы системы электропривода с трехтактным ЧИП.

Ключевые слова: частотно-импульсный преобразователь, инвертор, электропривод, моделирование

Цитирование: Дунаев М.П. Моделирование трехтактного частотно-импульсного преобразователя / М.П. Дунаев, А.А. Саушкин // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 94-101. – DOI: 10.25729/ESI.2024.35.3.008.

Введение. В современных системах управляемого электропривода постоянного тока [1] частотно-импульсный преобразователь (ЧИП) наряду с управляемым выпрямителем и широтно-импульсным преобразователем входят в первую тройку применяемых управляемых преобразователей постоянного тока [2].

Как правило, сложный технический объект, а именно к таким относятся объекты электропривода, можно разбить на некоторое количество взаимосвязанных блоков, каждый из которых имеет самостоятельное значение. В этом случае математическое описание технического объекта может быть представлено в виде функциональной модели. Функциональная модель является графическим изображением объекта, в котором каждый блок (функциональный элемент) обозначается прямоугольником. Входы и выходы функциональных элементов соответственно отмечены входящими и исходящими стрелками, показывающими направление прохождения сигналов.

При разработке функциональной модели трехтактного частотно-импульсного преобразователя учитывались общепринятые допущения [3 - 8].

Функциональная схема ЧИП [2] представлена на рис. 1, где обозначено: *ВА* – вводной автоматический выключатель, *СТ* – силовой трансформатор, *В* – силовой выпрямитель, *СФ* – сглаживающий фильтр, *СС* – силовая схема преобразователя, *Н* – нагрузка преобразователя, *БП* – блок питания системы управления, *ЗГ* – задающий генератор импульсов, *УО* – управляющий орган, *РИ* – распределитель импульсов по каналам управления, *УФ* – усилитель-формирователь импульсов управления, *ГР* – устройство гальванической развязки, $\sim U_c$ – напряжение питающей сети переменного тока, $U_{вх}$ – входное напряжение силовой схемы преобразователя, $U_{вых}$ – выходное напряжение преобразователя, $U_{бп}$ – выходное напряжение блока питания, U_y – выходное напряжение управляющего органа, U_3 – задающее напряжение.

В качестве математической модели ЧИП при исследовании динамических режимов его работы можно использовать передаточную функцию [2]

$$W_{ЧИП}(p) = \frac{U_{ВЫХ}(p)}{U_Y(p)} = K_{ЧИП} * e^{-p\tau} ,$$

где $U_{ВЫХ}$ – среднее выходное напряжение ЧИП; U_Y – напряжение управления ЧИП; $K_{ЧИП}$ – коэффициент усиления ЧИП; τ – время запаздывания реакции силовой части ЧИП на изменение сигнала управления U_Y .

Однако для более точного расчета характеристик ЧИП предпочтительнее использовать имитационное моделирование схемы ЧИП.

Моделирование схемы трехтактного частотно-импульсного преобразователя. Рассмотрим частотно-импульсный преобразователь с резонансным инвертором [9], схема которого изображена на рис. 2.

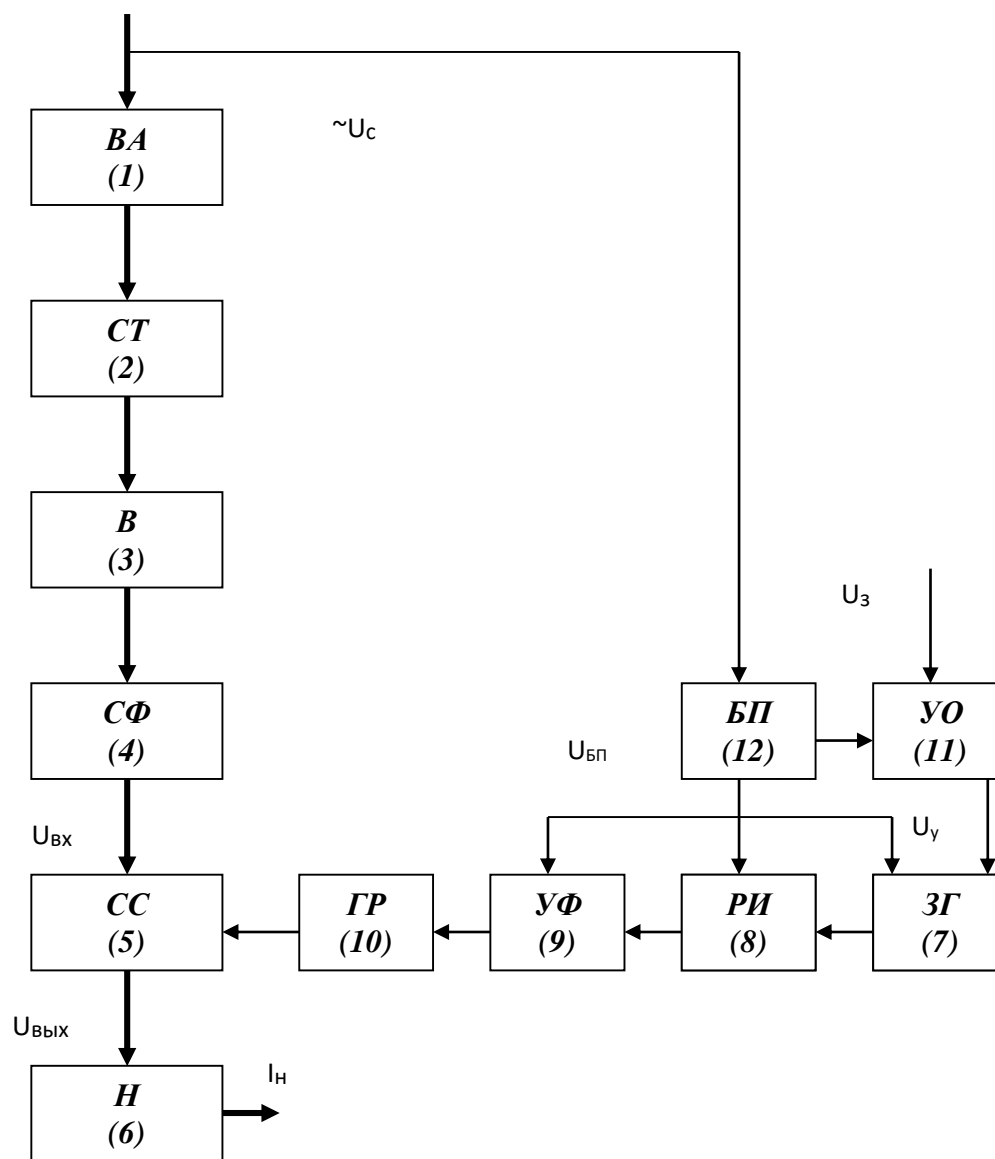


Рис. 1. Функциональная схема ЧИП

Резонансный инвертор имеет: трехфазный мост из тиристоров VS1, VS2, VS3, VS4, VS1, VS2; коммутирующие конденсаторы C1, C2, C3; коммутирующий дроссель L1; источник питания E; нагрузку Z1.

Инвертор работает в трехтактном режиме на нагрузку Z1. Для замыкания тока нагрузки во время закрытого состояния тиристоров она может быть шунтирована диодом VD1.

Устройство работает следующим образом. Попарно импульсами управления открываются тиристоры VS1, VS2; VS3, VS4; VS5, VS6 и т.д. Происходит перезаряд коммутирующих конденсаторов по колебательному закону в контуре, содержащем коммутирующий дроссель L1, нагрузку Z1, и соответствующую пару тиристоров (VS1, VS2 или VS3, VS4 или VS5, VS6).

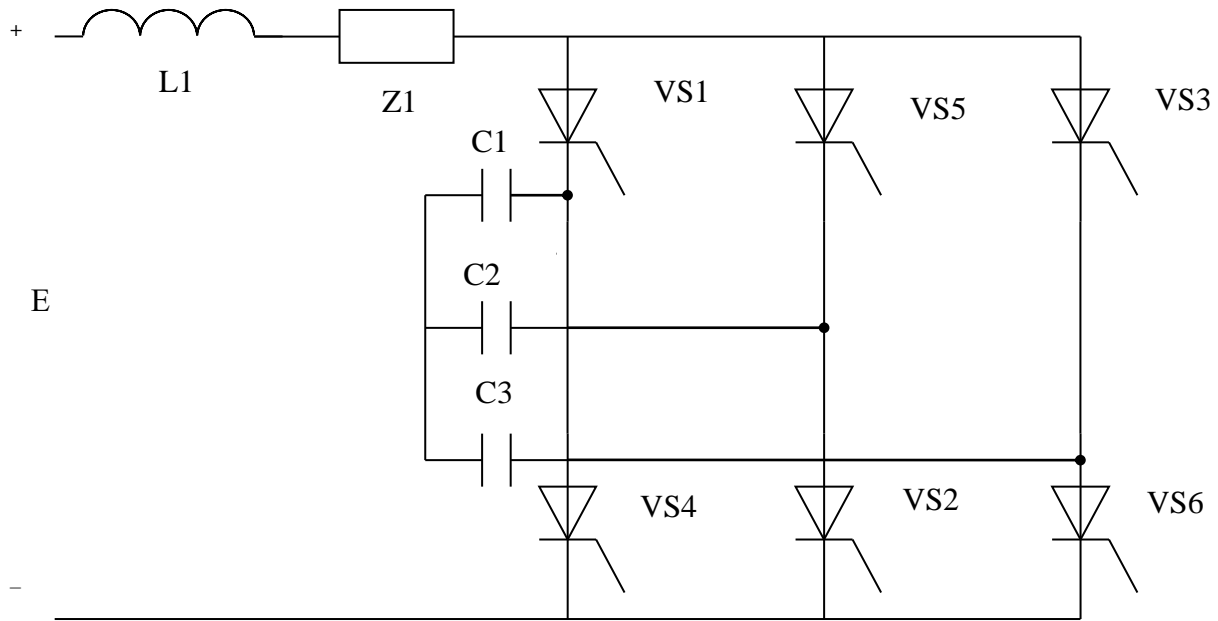


Рис. 2. ЧИП с резонансным инвертором

В установившемся режиме работы инвертора напряжение на коммутирующих конденсаторах является знакопеременным и периодическим во времени. В нагрузке Z1 протекает ток, частота которого втрое превышает частоту включения пар тиристоров.

На рис. 3 представлена имитационная модель трехфазного ЧИП в среде Matlab/Simulink [10-13], работающего на активную нагрузку.

С помощью модели, изображенной на рис. 3, были определены осциллограммы напряжений и токов на элементах виртуальной модели. Данные осциллограммы представлены на рис. 4.

Зависимость среднего напряжения на нагрузке $U_{н\text{ ср}}$ от частоты следования импульсов выходного напряжения ЧИП f (регулирующая характеристика ЧИП) описывается формулой [2]:

$$U_{н\text{ ср}} = S \cdot U_n = K_f \cdot U_n, \quad (1)$$

где S – скважность импульсов выходного напряжения ЧИП,

K_f – коэффициент передачи ЧИП,

U_n – напряжение источника питания ЧИП.

Уравнение электромеханической характеристики электропривода по системе ЧИП – Д в режиме непрерывных токов в относительных единицах будет иметь вид:

$$\Omega^* = S - I_{ср} / I_k, \quad (2)$$

где Ω^* – относительное значение частоты вращения двигателя, $\Omega^* = \Omega / \Omega_0$;

Ω – текущее значение частоты вращения двигателя,

Ω_0 – значение частоты вращения двигателя при идеальном холостом ходе, $\Omega_0 = U_n / c$;

$I_{ср}$ – средний ток якоря двигателя;

I_k – ток короткого замыкания, $I_k = U_n / (R_n + R_я)$;

R_n – внутреннее сопротивление ЧИП;

$R_я$ – сопротивление якоря двигателя.

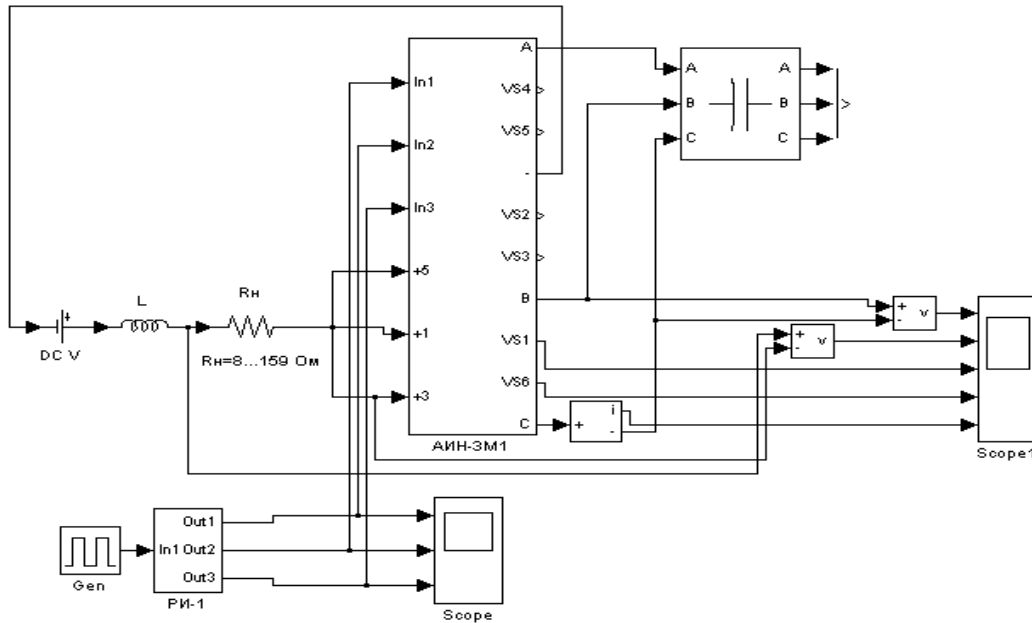


Рис. 3. Имитационная модель трехфазного ЧИП при работе на активную нагрузку

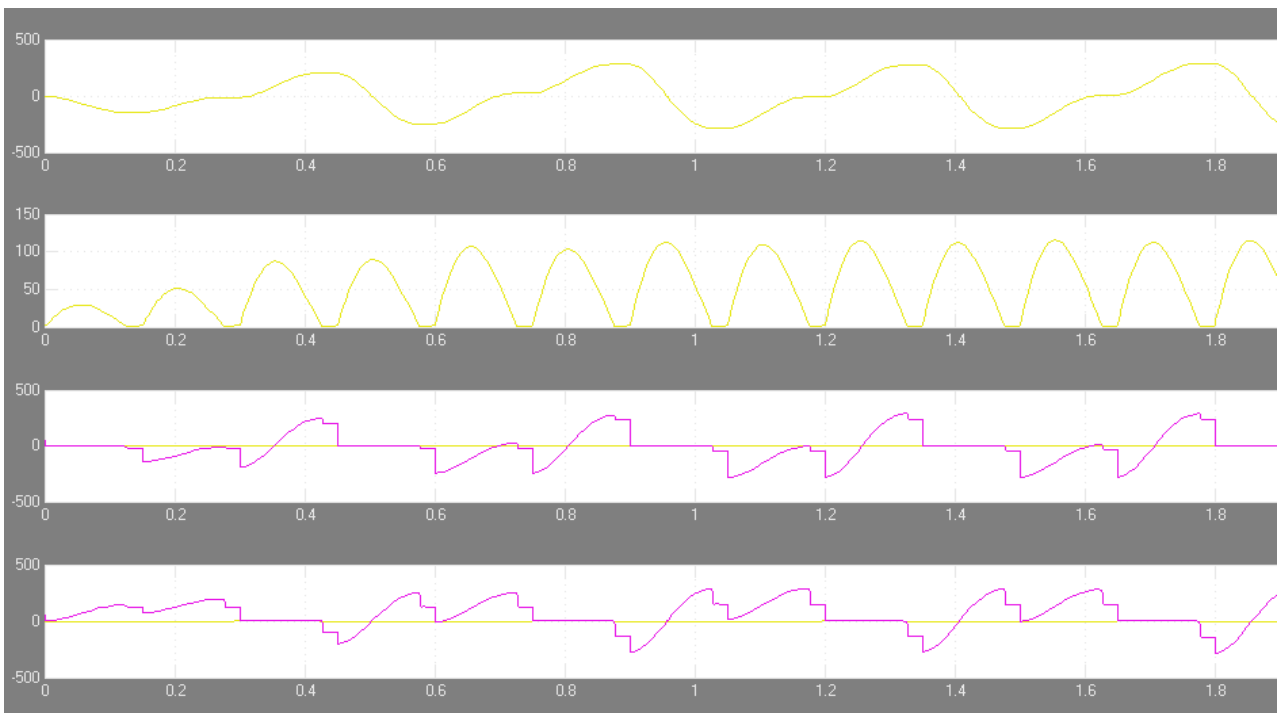


Рис. 4. Осциллограммы, снятые с датчиков модели трехфазного ЧИП при работе на активную нагрузку

Модель системы ЧИП-Д приведена на рис. 5 и предназначена для изучения статических и динамических характеристик электропривода с последовательным ЧИП с резонансной коммутацией.

В состав модели входят: источник постоянного напряжения (V1); силовая часть (тиристоры T1 – T6); задающий генератор (f); распределитель импульсов управления (R1 – R2); устройство задания нагрузки (L); осциллограф (Scope); двигатель постоянного тока с независимым возбуждением (DC Machine).

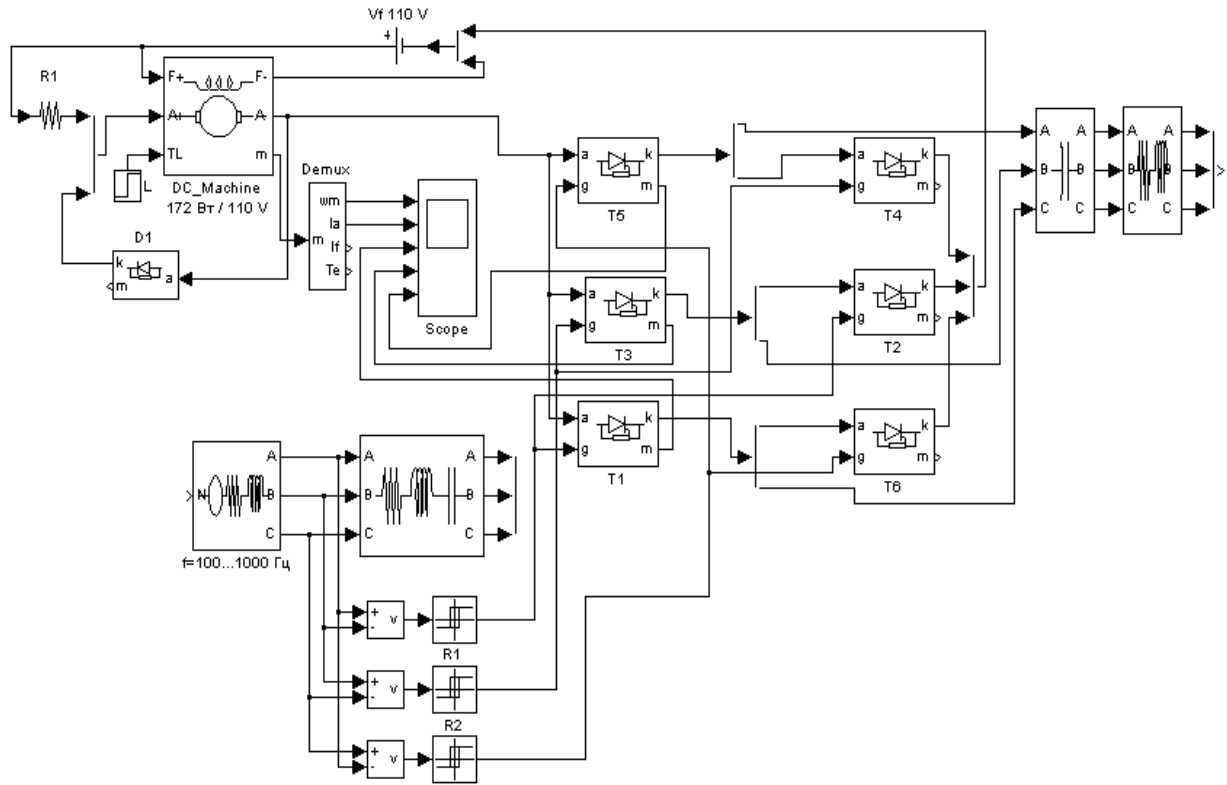


Рис. 5. Модель системы ЧИП-Д

Электромеханические характеристики системы ЧИП-Д представлены на рис. 6.

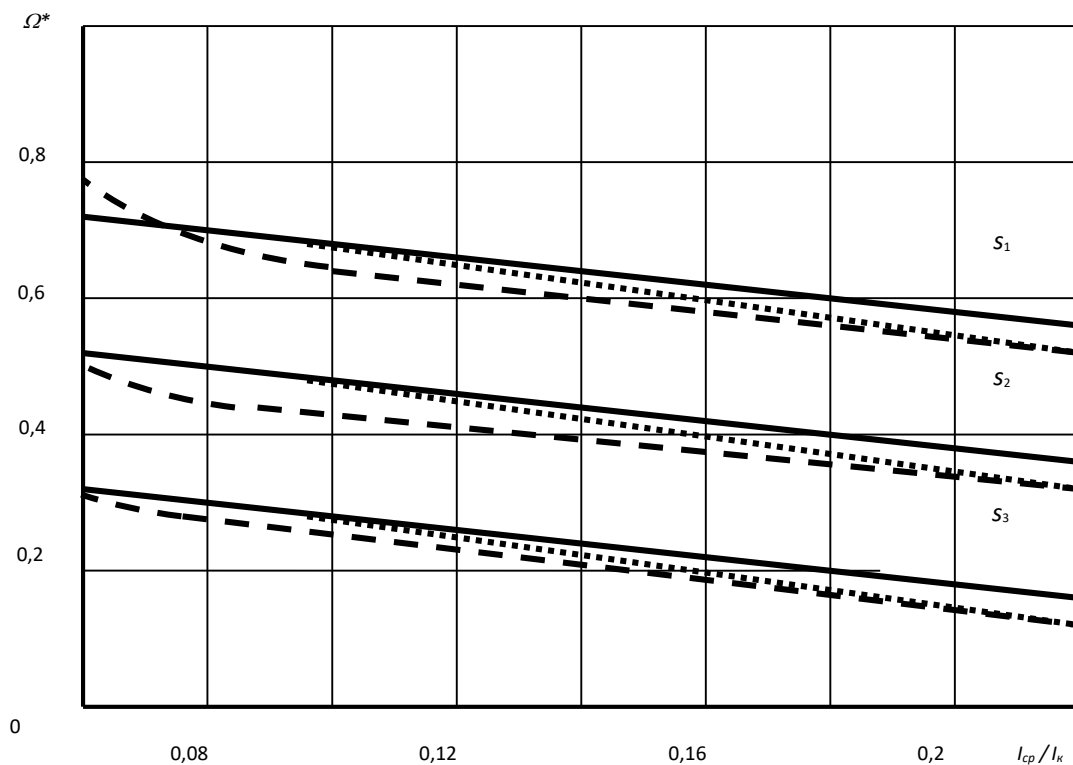


Рис. 6. Электромеханические характеристики системы ЧИП-Д

Расчетные электромеханические характеристики системы ЧИП-Д, построенные по выражению (2), показаны на рис. 6 сплошными линиями. Смоделированные электромеханические характеристики системы ЧИП-Д, полученные с помощью модели рис. 5, показаны на рис.

6 точками. Экспериментальные электромеханические характеристики системы ЧИП-Д [14] показаны на рис. 6 пунктиром. Параметры скважности на рис. 6 соответствуют следующим числовым значениям: $s_1=0,73$; $s_2=0,55$; $s_3=0,36$.

Характеристики рис. 6 представлены только в зоне непрерывных токов, так как из-за значительной частоты следования импульсов выходного напряжения ЧИП ($f_{\text{макс}}=900$ Гц) экспериментальное определение электромеханических характеристик системы ЧИП-Д в зоне прерывистых токов оказалось затруднительным.

Сравнивая представленные на рис. 6 расчетные, модельные и экспериментальные характеристики, можно увидеть, что погрешность электромеханических характеристик, полученных при помощи математического моделирования при режимах, близких к номинальному, не превышает 8%.

Таким образом, построенные в среде MATLAB математические модели системы ЧИП-Д с достаточной точностью описывают статические режимы ЧИП с резонансным инвертором.

Заключение. Представлена функциональная схема трехтактного частотно-импульсного преобразователя (ЧИП). Разработаны имитационные модели трехтактного ЧИП и системы электропривода ЧИП-Д в среде Matlab/Simulink. Проведено моделирование трехтактного ЧИП и системы электропривода ЧИП-Д. Приведено сравнение результатов моделирования и расчета электромеханических характеристик системы электропривода ЧИП-Д. Показано, что построенная в среде MATLAB математическая модель ЧИП-Д с достаточной точностью описывает статические режимы системы электропривода с трехтактным ЧИП.

Список источников

1. Кругликов О.В. Разработка, производство и применение энергоэффективных электрических машин и приводов / О.В. Кругликов, Л.Н. Макаров, С.В. Пискунов // Электротехника, 2015. – № 3. – С. 4-8.
2. Дунаев М.П. Резонансные инверторы для управления электроприводами / М.П. Дунаев. – Иркутск, Изд-во ИрГТУ, 2004. – 103 с.
3. Дунаев М.П. Моделирование потерь мощности в преобразователе частоты / М.П. Дунаев, С.У. Довудов // Электротехнические системы и комплексы, 2021. – № 2(51). – С. 45-51.
4. Дунаев М.П. Сравнение энергоэффективности схем автономных инверторов напряжения с синусоидальной широтно-импульсной и частотно-импульсной модуляцией / М.П. Дунаев, С.У. Довудов // Электротехнические системы и комплексы, 2021. – № 4(53). – С. 50-55.
5. Dunaev M.P., Dovudov S.U., Arshinskiy L.V. Energy characteristics the autonomous voltage inverter with the pulse-frequency modulation. International scientific conference “FarEastCon” 6-9 October 2020, Far Eastern federal university, Vladivostok.
6. Sharifov B., Gulyamov K., Yunusov R., Dovudov S. et al. The mathematical model of pulse width modulation frequency converter. Proceedings of the 2021 Ural-Siberian smart energy conference, USSEC 2021, Novosibirsk, 13–15 November 2021, Novosibirsk, 2021, pp. 81-84.
7. Dovudov S.U., Dunaev M.P., Zhiravetska A. et al. Analysis and comparison of energy indices of dc-dc pulse converters. Przegląd Elektrotechniczny, 2022, vol. 98, no 6, pp. 99-103.
8. Дунаев М.П. Анализ энергетических показателей импульсных преобразователей / М.П. Дунаев, С.У. Довудов // Вестник ИрГТУ, 2020. – №2. – С. 345-355.
9. Дунаев М.П., Марковец Ю.С. Однофазный последовательный инвертор / Патент на полезную модель №74018, БИ № 21, 2008.
10. Дьяконов В.П. MATLAB 6.5 SP1/7 + SIMULINK 5/6. Основы применения. – М.: СОЛОН-Пресс, 2005. – 800 с.
11. Дунаев М.П. Моделирование однофазного двухуровневого автономного инвертора напряжения с частотно-импульсной модуляцией / М.П. Дунаев, С.У. Довудов // Информационные и математические технологии в науке и управлении, 2020. – № 2(18). – С.134-143.
12. Dunaev M.P., Dovudov S.U., Arshinskiy L.V. Simulation of losses in autonomous inverter circuits with pulse-width and pulse-frequency modulation. Computational Technologies, 2022, vol. 27, no 5, pp. 5-13.
13. Марковец Ю.С. Виртуальная модель для исследования электропривода с последовательным ЧИП с резонансной коммутацией / Ю.С. Марковец / Вестник ИрГТУ, 2007. – №1-2. – С. 49-50.

14. Дунаев М.П. Характеристики системы электропривода постоянного тока с резонансным инвертором / М.П. Дунаев, Ю.С. Марковец / Повышение эффективности производства и использования электроэнергии в условиях Сибири / Мат. всеросс. науч.-практ. конф. Иркутск, ИрГТУ, 2005. – С.163-166.

Дунаев Михаил Павлович. Доктор технических наук, профессор Иркутского национального исследовательского технического университета, Author ID: 315898, ORCID: 0000-0002-1523-5553, mdunaev10@mail.ru, 664074, Россия, Иркутская обл., Иркутск, ул. Лермонтова, 83.

Саушкин Алексей Анатольевич. Аспирант Иркутского национального исследовательского технического университета, ORCID: 0000-0001-5600-4615, dsu_1991@mail.ru, 664074, Россия, Иркутская обл., Иркутск, ул. Лермонтова, 83.

UDC 621.314.3

DOI:10.25729/ESI.2024.35.3.008

Modeling of a three-stroke frequency-pulse converter

Mikhail P. Dunaev, Aleksei A. Saushkin

National research Irkutsk state technical university,
Russia, Irkutsk, *mdunaev10@mail.ru*

Abstract. The article presents a functional diagram of the tree-takt pulse-frequency converter (PFC). Imitation models of tree-takt pulse-frequency converter and PFC-DCM electric drive system in the Matlab/Simulink environment have been developed. Modeling of tree-takt pulse-frequency converter and electric drive system is given. A comparison of the results of modeling and calculation electromechanical characteristics of PFC-DCM electric drive system is given. It is shown that the mathematical model of CHIP-D constructed by the authors describes with sufficient accuracy the static modes of an electric drive system with a three-stroke CHIP.

Keywords: frequency-impulse converter, inverter, electric drive, simulation

References

1. Kruglikov O.V., Makarov L.N., Piskunov S.V. Razrabotka, proizvodstvo i primeneniye energoeffektivnykh elektricheskikh mashin i privodov [Development, production and application of energy efficient electrical machines and drives]. *Elektrotehnika* [Electrical Engineering], 2015, no. 3, pp. 4-8.
2. Dunaev M.P. Rezonansnyye invertory dlya upravleniya elektroprivodami [Resonant inverters controlling electric drives]. Irkutsk. IrGTU [Irkutsk state technical university], 2004, 103 p.
3. Dunaev M.P., Dovudov S.U. Modelirovaniye poter' moshchnosti v preobrazovatele chastity [Simulation of losses in frequency converter]. *Elektrotekhnicheskiye sistemy i komplekсы* [Electrical systems and complexes], 2021, no. 2(51), pp. 45-51.
4. Dunaev M.P., Dovudov S.U. Sravneniye energoeffektivnosti skhem avtonomnykh inverterov napryazheniya s sinusoidal'noy shirotno-impul'snoy i chastotno-impul'snoy modulyatsiyey [Comparison of Energy efficiency of autonomous voltage inverters with sine wave pulse width and pulse frequency modulation circuits]. *Elektrotekhnicheskiye sistemy i komplekсы* [Electrical systems and complexes], 2021, no. 4(53), pp. 50-55.
5. Dunaev M.P., Dovudov S.U., Arshinskiy L.V. Energy characteristics the autonomous voltage inverter with the pulse-frequency modulation. International scientific conference "FarEastCon" 6-9 October 2020, Far Eastern federal university, Vladivostok.
6. Sharifov B., Gulyamov K., Yunusov R., Dovudov S. et al. The mathematical model of pulse width modulation frequency converter. Proceedings of the 2021 Ural-Siberian smart energy conference, USSEC 2021, Novosibirsk, 13-15 November 2021, Novosibirsk, 2021, pp. 81-84.
7. Dovudov S.U., Dunaev M.P., Zhiravetska A. et al. Analysis and comparison of energy indices of dc-dc pulse converters. *Przeglad Elektrotechniczny*, 2022, vol. 98, no 6, pp. 99-103.
8. Dovudov S.U., Dunayev M.P. Analiz energeticheskikh pokazateley impul'snykh preobrazovateley [Analysis of energy indicators of pulse converters]. *Vestnik IrGTU* [ISTU Bulletin], 2020, vol. 24, no 2, pp. 345-355.
9. Dunayev M.P., Markovets U. Odnofaznii posledovatel'nii inverter [Single-phase series inverter]. Patent na poleznuy model [Patent for utility model], no. 74018, BI no. 21, 2008.
10. Diakonov V.P. MATLAB 6.5 SP1/7 + SIMULINK 5/6. Osnovi primeneniya [Basics of application]. M. SOLON-Press, 2005, 800 p.

11. Dunayev M.P., Dovudov S.U Modelirovaniye odnofaznogo dvukhurovnevoogo avtonomnogo invertora napryazheniya s chastotno-impul'snoy modulyatsiyey [Simulation of a single-phase two-level autonomous voltage inverter with pulse frequency modulation]. Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii [Information and mathematical technologies in science and management], 2020, no. 2 (18), pp. 134-143.
12. Dunaev M.P., Dovudov S.U., Arshinskiy L.V. Simulation of losses in autonomous inverter circuits with pulse-width and pulse-frequency modulation. Computational Technologies, 2022, vol. 27, no 5, pp. 5-13.
13. Markovets U.S. Virtualnaia model dlia issledovaniya elektroprivoda s posledovatelnim CHIP s rezonansnoi kommutatsiei. [Virtual model for studying an electric drive with a sequential CHIP with resonant switching] Vestnik IrGTU [Bulletin of Irkutsk state technical university], 2007, no. 1-2, pp. 49-50.
14. Dunaev M.P., Markovets U.S. Harakteristiki sistemi elektroprivoda postoiannogo toka s rezonansnim [Characteristics of a DC electric drive system with a resonant inverter]. Povishenie effektivnosti proizvodstva i ispolzovaniya elektroenergii v usloviakh Sibiri, Mat. Vserossiiskoi nauch.-prakt. Konf. Irkutsk, IrGTU [Increasing the efficiency of production and use of electric power in Siberia. Proc. All-Russian scientific-practical. conf. Irkutsk, IrSTU], 2005, pp.163-166.

Dunaev Mikhail Pavlovich. Doctor of technical sciences, professor, Irkutsk national research technical university, Author ID: 315898, ORCID: 0000-0002-1523-5553, mdunaev10@mail.ru, 664074, Russia, Irkutsk region, Irkutsk, Lermontov str., 83.

Saushkin Aleksei Anatolyevich. Graduate student, Irkutsk national research technical university, ORCID: 0000-0001-5600-4615, leha-34714@inbox.ru, 664074, Russia, Irkutsk region, Irkutsk, Lermontov str., 83.

Статья поступила в редакцию 21.06.2024; одобрена после рецензирования 16.07.2024; принята к публикации 10.10.2024.

The article was submitted 06/21/2024; approved after reviewing 07/16/2024; accepted for publication 10/10/2024.

Цифровая экономика и управление

УДК 004.75

DOI:10.25729/ESI.2024.35.3.009

Сравнительный анализ систем управления научными рабочими процессами

Воскобойников Михаил Леонтьевич, Феоктистов Александр Геннадьевич

Институт динамики систем и теории управления им. В.М. Матросова СО РАН,
Россия, Иркутск, *mikev1988@mail.ru*

Аннотация. Стремительное развитие параллельных и распределенных вычислительных систем, телекоммуникационных технологий и облачных платформ обеспечило возможность разработки и применения научных приложений с целью подготовки и проведения крупномасштабных экспериментов с большими массивами данных. Зачастую создаваемые приложения предполагают сложную схему решения задач, базирующуюся на интегрированном выполнении процессов передачи, обработки и анализа данных, ресурсоемких вычислений и принятия решений. При этом математическое и программное обеспечение приложений может реализовываться различными группами специалистов из разных организаций и ориентироваться на разнородные вычислительные ресурсы. Все это обуславливает необходимость использования развитых средств проектирования, реализации, развертывания и выполнения научных рабочих процессов в рамках единой распределенной вычислительной среды, интегрирующей в конечном счете алгоритмические знания, программно-аппаратное обеспечение, данные и разнообразные сервисы. В настоящее время в качестве таких средств, как правило, выступают системы управления рабочими процессами. В этой связи статья посвящена обсуждению текущего состояния известных систем управления рабочими процессами, а также рассмотрению проблем, связанных с научными рабочими процессами для различных вычислительных сред. Актуализируются проблемы разработки и применения подобных систем, которые в настоящее время не решены в полной мере. В частности, отмечается необходимость учета предметной специфики и обеспечения масштабирования вычислений, востребованность сервис-ориентированных приложений, эффективность эксплуатации гетерогенных сред, интегрирующих высокопроизводительные ресурсы пользователей, кластерные ресурсы центров коллективного пользования, Grid-систем и облачных платформ.

Ключевые слова: распределенные вычисления, научные рабочие процессы, системы управления рабочими процессами

Цитирование: Воскобойников М.Л. Сравнительный анализ систем управления научными рабочими процессами / М.Л. Воскобойников, А.Г. Феоктистов // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С.102-111. – DOI: 10.25729/ESI.2024.35.3.009.

Введение. В настоящее время решение крупномасштабных фундаментальных и прикладных задач зачастую осуществляется на основе распределенных научных приложений (РНП), характеризующихся модульной структурой их прикладного программного обеспечения (ПО), представленной в виде вычислительной модели, развитым системным ПО (совокупностью программ, обеспечивающих построение приложения, организацию вычислений с помощью его прикладных модулей, обработку данных и взаимодействие различных категорий пользователей с приложением) и ориентацией на решение определенного класса задач. В современных приложениях схема решения задачи реализуется научным рабочим процессом (НРП, англ., Scientific Workflow) – информационно-вычислительной структурой, отражающей бизнес-логику предметной области относительно данных, ПО (набора прикладных модулей) и вычислительных ресурсов в процессе проведения экспериментов.

Системы создания приложений на основе НРП относятся к классу систем управления рабочими процессами (англ., Workflow Management Systems – WMS). В общем случае WMS предоставляют программную инфраструктуру для настройки, выполнения и мониторинга определенной последовательности задач, организованных в виде рабочего процесса со сложными зависимостями между решаемыми задачами [1].

В процессе разработки НРП возникает необходимость решения следующих важных проблем: согласованного использования разнородных вычислительных ресурсов; учета специфики предметных областей решаемых задач; интеграции системного и прикладного ПО при подготовке и проведении экспериментов; стандартизации спецификаций объектов предметной области и вычислительной модели, форматов представления и протоколов передачи данных, доступа пользователей к подсистемам приложения; управления вычислениями в интегрированной среде посредством взаимодействия исполнительной подсистемы с системами мониторинга, локальными менеджерами ресурсов (ЛМР, англ., Local Resource Manager – LRM) в узлах среды, и метапланировщиками, работающими на уровне среды в целом; обеспечения масштабирования вычислений; поддержки технологии In-Memory Data Grid (IMDG) с целью ускорения обработки данных [2].

Необходимость учета предметной специфики и обеспечения масштабирования вычислений обуславливает переход от использования вычислительной среды общего назначения (Grid-системы, облачной платформы, ресурсов центров коллективного пользования и др.) к предметно-ориентированной вычислительной среде (ПОВС) с рациональным сочетанием возможностей включаемых в ее состав информационно-вычислительных ресурсов, потребностей и накладных затрат, обуславливаемых особенностями предметных областей для конкретных классов решаемых задач. При этом особое внимание научного сообщества уделяется развитию сервис-ориентированного подхода к организации и применению НРП.

Статья посвящена сравнительному анализу функциональных возможностей известных WMS относительно процессов разработки НРП в различных распределенных вычислительных средах.

Среда выполнения приложений. В общем случае рассматривается гетерогенная распределенная вычислительная среда (ГРВС), которая может интегрировать собственные кластерные ресурсы пользователей, а также кластерные ресурсы ЦКП, Grid и облачные платформы. Кластерные ресурсы для выполнения приложений могут объединяться и управляться на основе виртуализации или контейнеризации.

НРП [3] базируется на традиционном понятии пакета прикладных программ [4] и представляет собой комплекс взаимосвязанных прикладных программ и средств системного обеспечения (программных и языковых), предназначенный для автоматизации решения определенного класса задач в конкретной предметной области в ГРВС. В то же время НРП характеризуется следующими особенностями:

- ориентация на решение класса задач, требующих проведения расчетов с задействованием больших объемов вычислительных ресурсов (процессорного времени, оперативной памяти, дискового пространства и других характеристик);
- общая задача подразумевает параллельное решение множества ее слабо связанных или полностью независимых подзадач;
- алгоритмические знания (прикладное ПО) представляются в виде набора программных модулей, реализующих схемы решения задач;
- в процессе вычислений варианты данных параллельно обсчитываются экземплярами модулей;
- не предполагается интенсивного взаимодействия между параллельными вычислительными процессами;
- НРП представляет собой схему решения задачи;
- вычисления выполняются, как правило, по одной слабо меняющейся схеме, требующей динамического управления процессом вычислений;
- управление вычислениями осуществляет диспетчер НРП;

– необходимо использование связующего ПО.

Общая схема функционирования ГРВС для выполнения РППП приведена на рис. 1.

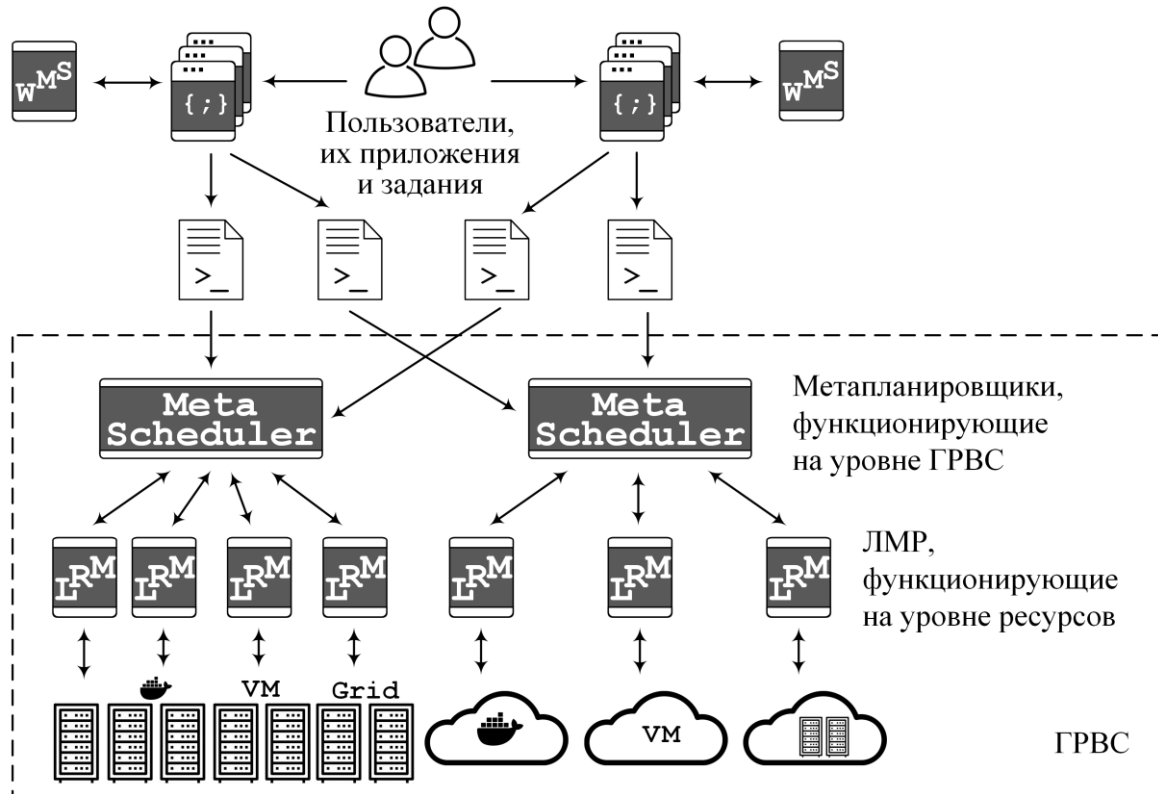


Рис. 1. Общая схема функционирования ГРВС

Пользователи приложений формируют задания в виде текстовых спецификаций вычислительных процессов, определяемых НРП. Такие спецификации включают сведения о требуемых ресурсах, исполняемом прикладном ПО, входных и выходных данных, а также иную необходимую дополнительную информацию. Например, дополнительная информация может определять критерии качества решения своей задачи: время, стоимость, надежность, безопасность и другие характеристики.

Основными компонентами ГРВС являются высокопроизводительные кластеры, кластеры виртуализированных и контейнеризированных ресурсов, которые используются в рамках таких инфраструктур, как ЦКП, кластерные Grid и облачные платформы. ГРВС может интегрировать компоненты перечисленных выше инфраструктур.

В общем случае управление вычислениями в ГРВС включает следующие уровни:

- уровень приложений, где выполняется планирование вычислений, построение НРП и выбор вычислительных ресурсов для его выполнения;
- уровень среды, на котором потоки заданий попадают в очередь метапланировщиков ГРВС, поддерживающих взаимодействие с выбранными ресурсами, и затем, в соответствии с дисциплинами диспетчеризации, им назначаются конкретные ресурсы;
- уровень ресурсов, где задания распределяются ЛМР между узлами с целью их дальнейшего выполнения.

Системы управления научными рабочими процессами. НРП представляется специализированной формой графа, которая описывает вычислительные процессы, их зависимости в рамках сбора, обработки и анализа данных. Он представляет собой бизнес-логику предметной области в применении предметно-ориентированных данных и ПО (набора прикладных модулей) для решения задач в этой предметной области.

Различают два типа НРП: абстрактный и конкретный. Абстрактный НРП описывается в виде абстрактной схемы решения задачи без обращения к конкретным ресурсам. Это позволяет определять НРП без детальной конкретизации его реализации на низком программно-аппаратном уровне. Операции абстрактного НРП могут реализовываться различным прикладным ПО и отображаться на любые подходящие ресурсы с помощью механизмов планирования вычислений и управления ресурсами. В случае конкретного НРП его операции связаны с предопределенным ПО, выполняемым на заданных ресурсах.

Построение НРП осуществляется по процедурной или непроцедурной постановкам задач. В первом случае задается последовательность операций НРП и затем осуществляется информационное планирование его входных и выходных параметров. Во втором случае, задача формулируется следующим образом: «вычислить значения целевых (выходных) параметров по заданным значениям исходных (входных) параметров». Далее на вычислительной модели (описании предметной области – ее параметров, вычислительных операций и операций обработки данных, а также отношений между параметрами и операциями) осуществляется автоматическое планирование последовательности операций НРП, необходимых для решения поставленной задачи. Как правило, планирование вычислений осуществляется одним из системных компонентов WMS. Непроцедурная постановка задачи позволяет разработчикам приложения и его пользователям работать на более высоком уровне абстракции. Они скрывают детали реализации и позволяют сосредоточиться на описании задачи и отношений между объектами.

WMS – это программное средство, частично автоматизирующее определение и выполнение НРП, а также управление им в соответствии с его информационно-вычислительной структурой. Для задания НРП часто используется ориентированный ациклический граф (DAG). В общем случае вершины и ребра DAG соответственно представляют прикладные программные модули и потоки данных между ними.

Следует отметить следующие WMS, которые хорошо известны, поддерживаются, развиваются и широко применяются на практике: Uniform Interface to Computing Resources (UNICORE) [5], Directed Acyclic Graph Manager (DAGMan) [6], Pegasus [7], Apache Airflow (AA) [8], HyperFlow [9], Workflow-as-a-Service Cloud Platform (WaaS CP) [10], Galaxy [11] и Orlando Tools (OT) [12]. UNICORE, DAGMan, Pegasus, AA и OT находятся на лидирующих позициях в области традиционного управления НРП. Они основаны на модульном подходе к созданию таких процессов. В рамках НРП общая задача разбивается на набор задач, обычно представленных в виде DAG. Как правило, UNICORE, DAGMan, Pegasus, AA, OT и другие подобные системы предоставляют свою собственную нестандартизированную модель НРП. Это затрудняет совместное использование спецификаций рабочих процессов и ограничивает совместимость приложений, разработанных с использованием различных WMS.

Комплексная поддержка веб-сервисов является актуальным направлением современных WMS. Использование веб-сервисов значительно расширяет вычислительные возможности приложений, основанных на НРП. WMS, ориентированные на работу с веб-сервисами, позволяют гибко и динамично интегрировать различные НРП от разных разработчиков за счет управления данными и вычислениями при выполнении этих процессов. В этом направлении успешно развиваются HyperFlow, WaaS CP и Galaxy.

В научных приложениях для решения различных классов задач в области мониторинга окружающей среды, особое значение имеет интеграция с геоинформационными системами посредством специализированных сервисов веб-обработки (WPS). Примерами проектов, направленных на создание и применение таких программных комплексов, являются GeoJModelBuilder (GJMB) [13] и Business Process Execution Language (BPEL) Designer Project

(DP) [14]. GJMB и BPEL DP относятся к классу WMS. GJMB – это фреймворк для управления и координации геопространственных датчиков, данных и прикладного ПО в среде, основанной на НРП. BPEL DP реализует интеграцию веб-сервисов и WPS с помощью рабочих процессов на основе BPEL. BPEL определяет модель для описания поведения сервис-ориентированных НРП в терминах взаимодействий (совокупности сообщений) между процессами и их партнерами (внешними сервисами). Другими преимуществами BPEL являются следующие:

- НРП могут не только вызывать веб-сервисы, но и сами быть представленными в виде сервисов;
- широкий спектр элементов управления и работы с данными, включающий элементы определения сложных структур данных и параллельных процессов их обработки, циклы, ветвления, подпроцессы, элементы реализации асинхронного взаимодействия веб-сервисов и др.;
- использование языка Web Service Description Language (WSDL) для описания интерфейсов веб-сервисов обеспечивает гибкую интеграцию с другими НРП и веб-приложениями;
- детальное описание НРП реализует оркестровку внутренних и внешних сущностей процесса, а спецификация процесса обмена сообщениями отражает хореографию внешних сущностей (вызываемых веб-сервисов).

Ключевые возможности WMS, которые наиболее важны с точки зрения разработки РНП на основе НРП, приведены в таблице 1.

Таблица 1. Возможности WMS для поддержки НРП

WMS	Характеристики процесса построения НРП							Особенности организации вычислений					
	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}
UNICORE	+ / - / - / - / - / -	+ / +	+	- / +	+ / -	+	-	+ / + / -	- / + / -	-	-	-	-
DAGMan	- / + / - / - / - / -	- / -	+	+ / -	+ / -	-	-	+ / + / -	+ / + / +	+	-	-	-
Pegasus	+ / - / - / - / - / -	- / -	+	+ / -	+ / -	-	-	+ / + / -	+ / + / +	+	-	-	-
AA	- / - / + / - / - / -	- / -	-	- / +	+ / -	-	-	+ / + / +	- / - / +	+	-	-	-
HyperFlow	- / - / - / - / + / +	- / -	+	+ / +	+ / -	+	-	+ / - / -	- / - / +	-	-	-	-
WaaS CP	+ / - / - / - / - / -	- / -	-	+ / -	+ / -	+	-	+ / - / -	- / + / +	+	-	-	-
Galaxy	+ / - / - / - / - / -	- / +	+	+ / +	+ / -	+	-	+ / - / -	- / - / +	-	-	-	-
OT	+ / - / - / - / - / -	+ / +	+	+ / +	+ / +	-	-	+ / + / +	+ / + / +	-	+	+	-
GJMB	+ / - / - / - / - / -	- / -	+	+ / +	+ / -	+	+	+ / - / -	+ / + / +	-	-	-	-
BPEL DP	- / - / - / + / - / -	+ / +	-	+ / +	+ / -	+	+	+ / - / -	+ / + / +	-	-	-	-

Рассматриваются следующие возможности WMS:

- c_1 – поддержка языков описания НРП: XML-like / Script-like / Python / BPEL / JavaScript / JSON;
- c_2 – поддержка управляющих языковых конструкций: ветвления / циклы;
- c_3 – поддержка возможности включения в НРП системных операций;
- c_4 – поддержка типов НРП: абстрактный / конкретный;
- c_5 – поддержка постановок задач: процедурная / непроцедурная;
- c_6 – поддержка вызова веб-сервисов в НРП;
- c_7 – поддержка стандарта WPS-сервисов;
- c_8 – уровень поддержки распараллеливания вычислений: задача / данные / конвейер;
- c_9 – поддержка следующих типов вычислительной среды: кластер / Grid / облачная среда;
- c_{10} – необходимость подключения дополнительного связующего ПО;
- c_{11} – поддержка технологии IMDG;

- c_{12} – автоматизация непрерывной интеграции ПО;
- c_{13} – автоматизация контейнеризации ПО.

WMS используют XML-подобные или скриптовые языки для описания предметных областей приложений и самих НПП. BPEL характеризуется рядом преимуществ для спецификации НПП, перечисленных выше. Однако его языковых средств зачастую бывает недостаточно, чтобы описать сложную вычислительную модель. Поэтому целесообразно применять комбинацию XML-подобного языка и BPEL. Ветвления и циклы в рабочем процессе позволяют при необходимости более гибко организовать вычисления. UNICORE, OT и BPEL DP предоставляют такую поддержку в полном объеме для построения НПП, не связанного с DAG. Дополнительным преимуществом UNICORE, DAGMan, Pegasus, HyperFlow, Galaxy, GJMB и OT является возможность включения системных операций в структуру рабочего процесса. К таким операциям относятся пред- и постобработка данных, мониторинг среды, взаимодействие с компонентами локальных менеджеров ресурсов и т.д.

HyperFlow, WaaSCP, Galaxy, OT, GJMB и BPEL DP позволяют строить как абстрактные, так и конкретные НПП. При этом OT обеспечивает оба вида постановок задач, на основе которых выполняется построение НПП: процедурную и непроцедурную.

HyperFlow, WaaSCP, Galaxy, GJMB и BPEL DP ориентированы на работу с веб-сервисами. Кроме того, GJMB и BPEL DP поддерживают стандарт WPS.

Большая часть рассматриваемых WMS являются системами общего назначения и могут успешно применяться в различных предметных областях. Некоторые системы являются более специализированными. Например, Galaxy ориентирована на поддержку проведения геномных исследований, а GJMB и BPEL DP позволяют решать задачи в области геоинформатики.

Уровень задач означает, что задачи, определенные структурой рабочего процесса, выполняются на независимых узлах. На уровне данных набор данных делится на подмножества. Каждое подмножество обрабатывается на отдельном узле всеми или частью прикладных программных модулей, включенных в рабочий процесс. На уровне конвейера последовательное выполнение прикладных модулей для обработки данных выполняется одновременно на разных подмножествах набора данных. AA и OT предоставляют все уровни распараллеливания вычислений. DAGMan, Pegasus, OT, GJMB и BPEL DP допускают использование всех типов вычислительной среды. Однако системе Pegasus требуется интеграция с DAGMan для управления заданиями приложений на уровне среды. OT предоставляет системные операции для поддержки технологии IMDG на основе связующего ПО Apache Ignite. Инструменты автоматизации непрерывной интеграции прикладного и системного ПО предоставляются в OT. В то же время обеспечение автоматизации процессов контейнеризации ПО является не решенной в полной мере проблемой для всех известных WMS. Практически во всех WMS слабо проработаны вопросы моделирования вычислительных процессов. Решение упомянутых вопросов требует включения в WMS дополнительных средств взаимодействия с контрольно-измерительными системами, профилировщиками ПО, инструментами анализа данных и моделирования работы как отдельных информационно-вычислительных компонентов ГРВС, так и всей среды в целом.

К сожалению, из-за разницы в функциональных и системных возможностях WMS затруднительно провести адекватное сравнение трудозатрат их пользователей на разработку приложений, подготовку и проведение вычислительных экспериментов, а также оценить эффективность использования вычислительных ресурсов при решении практических задач. Однако некоторые результаты сравнительного анализа управления выполнением ряда известных рабочих процессов метапланировщиками на уровне среды можно найти в [15].

Направления развития WMS. Вопросы разработки и применения НРП в разных областях исследований требуют постоянного их изучения, сравнения и развития с целью поддержки междисциплинарного сотрудничества. Потребность совместного использования НРП, миграции этих рабочих процессов между различными вычислительными средами, их адаптации к различным категориям пользователей, языкам и программным средствам обоснованно обуславливают необходимость совершенствования WMS. На основе проведенного сравнительного анализа можно определить направления развития функциональных возможностей WMS. В том числе выделить следующие важные аспекты дополнительных разработок для таких систем:

- предоставление дружественного пользовательского интерфейса для взаимодействия с компонентами инструментального комплекса и создаваемых сервис-ориентированных приложений разным категориям пользователей (разработчикам, администраторам и пользователям), поддерживающего как текстовые языки представления структурированных данных, так и языки для ввода информации с помощью веб-форм с последующим автоматическим конвертированием полученных спецификаций на соответствующие текстовые языки;
- корректное описание вычислительной модели предметной области (параметров, операций, логических выражений, продукций и модулей, вычислительных ресурсов, элементов вычислительной истории, характеристик административных политик управления ресурсами и др., а также отношений между перечисленными объектами);
- работа со сложными структурами данных, такими, как составные параметры и параллельные списки данных, а также неструктурированными файловыми данными;
- статическое планирование вычислений по непроцедурной постановке задачи, построение НРП по процедурной постановке задачи с использованием различных операторов ветвления и циклов, а также статико-динамическое планирование расчетных схем и распределение возникающей нагрузки на ресурсы при работе с параллельными списками данных;
- обеспечение средств создания и применения сервис-ориентированных НРП при сохранении поддержки традиционных НРП с исполняемыми модулями;
- возможность включения в состав НРП системных операторов обработки и анализа данных, статико-динамического планирования, конвейеризации вычислений, параллельного выполнения алгоритмов;
- генерация НРП на языке программирования общего назначения с целью их последующего автономного запуска;
- автоматизация непрерывной интеграции и контейнеризации ПО, включая моделирование вычислительных процессов и тестирование ПО на ресурсах ГРВС.

Заключение. Разработка научных приложений, основанных на рабочих процессах, является перспективной парадигмой для решения широкого спектра фундаментальных и прикладных задач в различных областях знаний. Использование высокопроизводительных вычислительных систем обеспечивает высокую эффективность процесса решения. В этой связи научное сообщество прилагает значительные усилия в развитии данной парадигмы. В частности, пристальное внимание уделяется системам управления НРП.

В рамках проведенного исследования рассмотрены основные понятия, связанные с НРП и системами управления ими. Проведен сравнительный анализ известных систем управления НРП. Оценены тенденции современного развития РНП на основе сервис-ориентированных НРП. Обсуждены вопросы и проблемы исследований в этой области, не решенные в полной мере, в том числе определены актуальные направления развития систем управления НРП.

Дальнейшие исследования связаны с разработкой нового инструментального комплекса, развивающего функциональные возможности ОТ применительно к сервис-ориентированным приложениям, включая представление НРП на языке WPEL и поддержку работы с WPS-сервисами, а также автоматизацию процессов генерации программ на языках программирования общего назначения на основе НРП и контейнеризации прикладного и системного ПО в ГРВС. Реализация перечисленных выше средств позволит успешно использовать инструментальный комплекс для построения цифровых двойников компонентов информационно-вычислительных систем и их программно-аппаратных инфраструктур на основе системных НРП, дополнительно направленных на анализ и моделирование характеристик, свойств, состояний и процессов вычислительной среды.

Благодарности. Исследование проведено в рамках проекта № FWEW-2021-0005 «Технологии разработки и анализа предметно-ориентированных интеллектуальных систем группового управления в недетерминированных распределенных средах» (рег. № 121032400051-9) при поддержке Министерства науки и высшего образования РФ.

Список источников

1. Schael T. Workflow management systems for process organisations. Lecture notes in computer science, 1996, vol. 1096, p. 208, DOI:10.1007/978-3-662-21574-6.
2. Zhang H., Chen G., Ooi B.C. In-memory big data management and processing: A survey. IEEE Transactions on knowledge and data engineering, 2015, vol. 27, no. 7, pp. 1920–1948, DOI:10.1109/TKDE.2015.2427795.
3. Феоктистов А.Г. Автоматизация разработки и применения распределенных пакетов прикладных программ / А.Г. Феоктистов, И.А. Сидоров, С.А. Горский // Проблемы информатики, 2017. – № 4. – С. 61-78.
4. Горбунов–Посадов М.М. Системное обеспечение пакетов прикладных программ / М.М. Горбунов–Посадов, Д.А. Корягин, В.В. Мартынюк. – М.: Наука, 1990. – 208 с.
5. UNICORE. Available at: <https://www.unicore.eu/> (accessed: 09.04.2024).
6. DAGMan. Available at: <https://htcondor.org/dagman/dagman.html> (accessed: 04/09/2024).
7. Deelman E., Singh G., Su M.H. et al. Pegasus: A framework for mapping complex scientific workflows onto distributed systems. Scientific programming, 2005, vol. 13, pp. 219–237, DOI:10.1155/2005/128026.
8. Singh P. Learn PySpark: Build Python-based machine learning and deep learning models. Apress, 2019, pp. 67–84.
9. Balis B. HyperFlow: A model of computation, programming approach and enactment engine for complex distributed workflows. Future generation computer systems, 2016, vol. 55, pp. 147–162. DOI:10.1016/j.future.2015.08.015.
10. Hilman M.H., Rodriguez M.A., Buyya R. Workflow-as-a-service cloud platform and deployment of bioinformatics workflow applications. Knowledge management in the development of data-intensive systems, 2021, pp. 205–226, DOI:10.1201/9781003001188-14.
11. Jalili V., Afgan E., Gu Q. et al. The galaxy platform for accessible, reproducible and collaborative biomedical analyses. Nucleic acids research, 2020, vol. 48, no. W1, pp. W395-W402, DOI:10.1093/nar/gkaa434.
12. Gorsky S.A. Continuous integration, delivery, and deployment for scientific workflows in Orlando Tools. Proceedings of the 2nd international workshop on information, computation, and control systems for distributed environments, CEUR-WS Proceedings, 2020, vol. 2638, pp. 118-128, DOI:10.47350/ICCS-DE.2020.11.
13. Yue P., Zhang M., Tan Z. A geoprocessing workflow system for environmental monitoring and integrated modelling. Environmental modelling and software, 2015, vol. 69, pp. 128–140, DOI:10.1016/j.envsoft.2015.03.017.
14. Tan X., Jiao J., Chen N. et al. Geoscience model service integrated workflow for rain-storm waterlogging analysis. International journal of digital Earth, 2021, vol. 14, pp. 851–873, DOI:10.1080/17538947.2021.1898686.
15. Feoktistov A. Tender of computational works in heterogeneous distributed environment. CEUR-WS Proceedings, 2020, vol. 2638, pp. 99–108, DOI:10.47350/ICCS-DE.2020.09.

Воскобойников Михаил Леонтьевич. Младший научный сотрудник института динамики систем и теории управления им. В.М. Матросова СО РАН. Область научных интересов – распределенные вычисления, системы управления рабочими процессами, автоматизация разработки сервис-ориентированных приложений. AuthorID: 1102663, SPIN: 3417-0258, ORCID: 0000-0003-3034-4907, mikev1988@mail.ru, Россия, г. Иркутск, Лермонтова, 134.

Феоктистов Александр Геннадьевич. Д.т.н., доцент, главный научный сотрудник института динамики систем и теории управления им. В.М. Матросова СО РАН. Область научных интересов – распределенные вычисления, системы управления рабочими процессами, автоматизация разработки научных приложений, концептуальное моделирование. AuthorID: 1535, SPIN: 5743-1777, ORCID: 0000-0002-9127-6162, agf65@yandex.ru, Россия, г. Иркутск, Лермонтова, 134.

UDC 004.75

DOI:10.25729/ESI.2024.35.3.009

Comparative analysis of scientific workflow management systems

Mikhail L. Voskoboinikov, Alexander G. Feoktistov

Matrosov Institute for System Dynamics and Control Theory of SB RAS,
Russia, Irkutsk, *mikev1988@mail.ru*

Abstract. The rapid evolution of parallel and distributed computing systems, telecommunication technologies, and cloud platforms has enabled the development and use of scientific applications to prepare and conduct large-scale experiments with large amounts of data. Often, the applications implement a complex problem-solving scheme based on the integrated execution of processes for data transfer, processing and analysis, resource-intensive computation, and decision-making. At the same time, the mathematical models and software of applications may be developed by different groups of specialists from different organizations and focused on heterogeneous computing resources. This requires the use of advanced tools for the design, implementation, deployment, and execution of scientific workflows within a single distributed computing environment, ultimately integrating algorithmic knowledge, software and hardware used, data, and various services. Today, such tools are usually workflow management systems. In this context, the paper is dedicated to discuss the current state of known workflow management systems, as well as to address the problems associated with the development and use of scientific workflows in different computing environments. The problems associated with the development and use of such systems, which are currently not fully solved, are highlighted. In particular, we point out the need to take into account subject domain specificities, the computation scaling, the demand for service-oriented applications, and the efficiency of using heterogeneous distributed environments that integrate high-performance user resources, cluster resources of shared use centers, Grid systems, and cloud platforms.

Keywords: distributed computing, scientific workflows, workflow management systems

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. Schael T. Workflow management systems for process organisations. Lecture notes in computer science, 1996, vol. 1096, p. 208, DOI:10.1007/978-3-662-21574-6.
2. Zhang H., Chen G., Ooi B.C. In-memory big data management and processing: A survey. IEEE Transactions on knowledge and data engineering, 2015, vol. 27, no. 7, pp. 1920–1948, DOI:10.1109/TKDE.2015.2427795.
3. Feoktistov A.G., Sidorov I.F., Gorsky S.A. Avtomatizaciia razrabotki i primeneniia raspredelennykh paketov prikladnykh programm [Automation of development and application of distributed applied software packages]. Problemy informatiki [Problems of Informatics], 2017, no. 4, pp. 61-78.
4. Gorbunov–Posadov M. M., Koryagin D.A., Martynyuk V.V. Sistemnoe obespechenie paketov prikladnykh programm [System support of applied software packages]. Moscow, Nauka, 1991, 208 p.
5. UNICORE. Available at: <https://www.unicore.eu/> (accessed: 09.04.2024).
6. DAGMan. Available at: <https://htcondor.org/dagman/dagman.html> (accessed: 04/09/2024).
7. Deelman E., Singh G., Su M.H. et al. Pegasus: A framework for mapping complex scientific workflows onto distributed systems. Scientific programming, 2005, vol. 13, pp. 219–237, DOI:10.1155/2005/128026.
8. Singh P. Learn PySpark: Build Python-based machine learning and deep learning models. Apress, 2019, pp. 67–84.

9. Balis B. HyperFlow: A model of computation, programming approach and enactment engine for complex distributed workflows. *Future generation computer systems*, 2016, vol. 55, pp. 147–162. DOI:10.1016/j.future.2015.08.015.
10. Hilman M.H., Rodriguez M.A., Buyya R. Workflow-as-a-service cloud platform and deployment of bioinformatics workflow applications. *Knowledge management in the development of data-intensive systems*, 2021, pp. 205–226, DOI:10.1201/9781003001188-14.
11. Jalili V., Afgan E., Gu Q. et al. The galaxy platform for accessible, reproducible and collaborative biomedical analyses. *Nucleic acids research*, 2020, vol. 48, no. W1, pp. W395–W402, DOI:10.1093/nar/gkaa434.
12. Gorsky S.A. Continuous integration, delivery, and deployment for scientific workflows in Orlando Tools. *Proceedings of the 2nd international workshop on information, computation, and control systems for distributed environments*, CEUR-WS Proceedings, 2020, vol. 2638, pp. 118–128, DOI:10.47350/ICCS-DE.2020.11.
13. Yue P., Zhang M., Tan Z. A geoprocessing workflow system for environmental monitoring and integrated modelling. *Environmental modelling and software*, 2015, vol. 69, pp. 128–140, DOI:10.1016/j.envsoft.2015.03.017.
14. Tan X., Jiao J., Chen N. et al. Geoscience model service integrated workflow for rain-storm waterlogging analysis. *International journal of digital Earth*, 2021, vol. 14, pp. 851–873, DOI:10.1080/17538947.2021.1898686.
15. Feoktistov A. Tender of computational works in heterogeneous distributed environment. *CEUR-WS Proceedings*, 2020, vol. 2638, pp. 99–108, DOI:10.47350/ICCS-DE.2020.09.

Voskoboinikov Mikhail Leontevich. *Junior Researcher at the Matrosov institute for system dynamics and control theory of SB RAS. The main direction of research – distributed computing, workflow management systems, automation of developing service-oriented applications. AuthorID: 1102663, SPIN: 3417-0258, ORCID: 0000-0003-3034-4907, mikev1988@mail.ru, Russia, Irkutsk, Lermontova, 134.*

Feoktistov Alexander Gennadevich. *Doctor of technical sciences, associate professor, chief researcher at the Matrosov institute for system dynamics and control theory of SB RAS. The main direction of research – distributed computing, workflow management systems, automation of developing scientific applications, conceptual modeling. AuthorID: 1535, SPIN: 5743-1777, ORCID: 0000-0002-9127-6162, agf65@yandex.ru, Russia, Irkutsk, Lermontova, 134.*

Статья поступила в редакцию 10.04.2024; одобрена после рецензирования 19.09.2024; принята к публикации 08.10.2024.

The article was submitted 04/10/2024; approved after reviewing 09/19/2024; accepted for publication 10/08/2024.

УДК 519.237.5

DOI:10.25729/ESI.2024.35.3.010

Влияние вероятностных моделей работ, связанных с защитой информации, на значения показателей эффективности

Краковский Юрий Мечеславович, Киргизбаев Владислав Павлович

Иркутский государственный университет путей сообщения,

Россия, Иркутск, 79149267772@yandex.ru

Аннотация. Для повышения эффективности выполнения работ по защите информации в организации предложено использовать бюджетный фонд, в котором накапливаются, а затем расходуются необходимые финансовые средства. В условиях неопределенности интервалы времени по расходам и сами расходы являются случайными величинами, которые при моделировании нужно описывать вероятностными моделями. Проведено исследование о влиянии видов вероятностных моделей и значений их числовых характеристик на показатели эффективности работ, выполняемых сотрудниками службы информационной безопасности организации, обслуживающих корпоративную информационную систему. Показателями эффективности являются вероятность «обнуления» бюджетного фонда и коэффициент вариации, которые при дискретно-имитационном моделировании заменяются точечными и интервальными оценками. Получены практические рекомендации.

Ключевые слова: дискретно-имитационное моделирование, бюджетный фонд, экономика информационной безопасности, показатели эффективности

Цитирование: Краковский Ю.М. Влияние вероятностных моделей работ, связанных с защитой информации, на значения показателей эффективности / Ю.М. Краковский, В.П. Киргизбаев // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 112-119. – DOI: 10.25729/ESI.2024.35.3.010.

Введение. В связи с повышением роли информационных технологий и искусственного интеллекта в развитии мировой экономики в нашей стране также уделяется большое внимание «цифровизации экономики». В первую очередь выделим стратегию развития информационного общества в РФ на 2017–2030 годы и реализуемую в настоящее время программу «Цифровая экономика РФ». Реализация этой программы предъявляет повышенные требования к методам защиты информации, а также к роли экономики информационной безопасности.

Актуальность роли экономики информационной безопасности приведена в статье [1], где отмечено, что управление экономикой информационной безопасности включает в себя особые процессы и методы для эффективного планирования, управления, мониторинга и контроля событий, связанных с защитой информации. Особое внимание уделяется квалификации специалистов, поскольку она значительно влияет на стоимость управления экономикой информационной безопасности. Авторы статьи акцентируют внимание на динамике внешней инфраструктуры, связанной с научно-техническим прогрессом и увеличением угроз информационной безопасности, а также на изменениях в законодательстве, включая переход российских компаний на отечественное программное обеспечение. Отметим и другие работы, подтверждающие актуальность подобных исследований [2-4].

В статье авторов [4] приведено математическое описание моделирования работ, выполняемых сотрудниками службы информационной безопасности организации, на основе бюджетного фонда (БФ). Этот фонд выполняет две функции: а) по мере необходимости производится оплата работ, связанных с устранением инцидентов, когда для каждого вида работ определяется периодичность использования фонда (сут.) и стоимость этих работ (тыс. руб.); б) осуществляется накопление платежей с какой-то периодичностью (сут.) и величиной стоимости этих платежей (тыс. руб.). Поступления в фонд предлагается описывать не случайными величинами, как для интервалов, так и для размеров поступлений.

В условиях неопределенности интервалы времени между инцидентами и затраты для их устранения являются случайными величинами с известными вероятностными моделями в виде двухпараметрических функций распределений. Для определения значений параметров этих функций, которые необходимы при моделировании значений этих величин, требуется знать значения двух числовых характеристик:

- 1) математических ожиданий для интервалов (mt_j), сут. и для затрат (mz_j), тыс. руб.;
- 2) коэффициентов вариации для интервалов (kvt_j) и для затрат (kvz_j).

Здесь j – номер работы.

Целью данной статьи является исследование влияния видов вероятностных моделей и значений их числовых характеристик на значения показателей эффективности, характеризующих работу специалистов службы информационной безопасности организации, обслуживающих корпоративную информационную систему с целью ее защиты.

Математическое описание исследования. Данное исследование является продолжением работы [4]. Состояние БФ описывается случайным нестационарным процессом вида

$$Fs(t) = Fs_0 + \sum_{j=1}^m Y_j(t) - \sum_{j=1}^m Z_j(t), \text{ тыс. руб.}, \quad (1)$$

где Fs_0 – начальное значение процесса $Fs(t)$, тыс. руб.; $Y_j(t)$ – суммарная величина доходов для j -й работы за время t , тыс. руб.; $j = 1, 2, \dots, m$, m – число видов работ по устранению инцидентов информационной безопасности; $Z_j(t)$ – суммарная величина расходов для j -й работы за время t , тыс. руб.

Величину Fs_0 предлагается задавать в долях от величины средних расходов

$$Fs_0 = g \cdot X, \text{ тыс. руб.}, \quad (2)$$

где X – средние финансовые средства (тыс. руб.), необходимые для выполнения годового объема всех работ (расходы); g – коэффициент, равный 0,05 или 0,10.

Суммарная величина доходов для j -го вида работ за время t равна

$$Y_j(t) = Y_{oj} \cdot N_{oj}(t), \text{ тыс. руб.}, \quad (3)$$

где $N_{oj}(t)$ – число платежей за время t для j -го вида работ; Y_{oj} – значение платежа для j -го вида работ при пополнении БФ, тыс. руб. Таким образом доходы БФ создаются за счет поступления платежей от организации, которая является владельцем корпоративной информационной системы.

Суммарная величина расходов для j -го вида работ равна

$$Z_j(t) = \sum_{q=1}^{N_j(t)} z_{qj}, \text{ тыс. руб.}, \quad (4)$$

где z_{qj} – величина q -го расхода для j -й работы; $N_j(t)$ – число этих расходов за время t . Таким образом, расходы БФ создаются за счет оплаты работ по устранению инцидентов информационной безопасности.

Средняя суммарная величина доходов за год по каждой работе должна быть равна средней суммарной величине расходов за год по этой работе. Из этого условия определяются величины Y_{oj} . Подробнее это изложено в работе [4].

Вычисление показателей эффективности в зависимости от исходных данных реализуется моделирующей программой, использующей дискретно-имитационный подход [5, 6].

При моделировании процесса $Fs(t)$ (1) с учетом (3) и (4) по каждой реализации за время Tm определяются величины времени (s), когда первый раз $Fs(t) < 0$,

$$s = \min_t (t : Fs(t) < 0), \text{ сут.} \quad (5)$$

Эти величины преобразуются в упорядоченную по возрастанию этих значений выборку

$$Ts = (s_1, \dots, s_i, \dots, s_n), \quad (6)$$

обработка которой позволит найти выбранные показатели эффективности.

Учитывая нестационарность процесса (1), величины (5) существуют не для всех реализаций. Чтобы повысить точность имитационного моделирования в данном исследовании, число реализаций равно 20000.

Предлагаются два показателя эффективности:

1) вероятность «обнуления» БФ

$$p_t = P(s < S_t), \quad (7)$$

где S_t – интервал времени, через который может произойти «обнуление» БФ; число выборочных значений на этом интервале равно k_t ;

2) коэффициент вариации в процентах за время S_t

$$c_v = \frac{\sigma}{\mu} \cdot 100, (\%), \quad (8)$$

где σ – значение среднеквадратического отклонения на интервале S_t , μ – значение математического ожидания на интервале S_t .

При имитационном моделировании величины (7) и (8) оцениваются через точечные и интервальные оценки, полученные по выборке (6) за время S_t :

- 1) \tilde{R}_t – точечная оценка величины (7);
- 2) (τ_1, τ_2) – интервальная оценка величины (7);
- 3) \tilde{c}_v – точечная оценка величины (8);
- 4) (v_1, v_2) – интервальная оценка величины (8).

Для исследования используются следующие работы, предложенные совместно с экспертами:

- 1) поддержка и модернизация программных средств защиты информации;
- 2) восстановление работоспособности технических и программно-аппаратных средств защиты информации;
- 3) резервное копирование важной информации (облачное хранение, зеркалирование и т.д.);
- 4) поддержка, восстановление и модернизация средств защиты информации для сложных сетевых инфраструктур (Dallas Lock и т. д.);
- 5) поддержка и модернизация криптографических средств защиты информации, включая программно-аппаратные комплексы.

Данные работы соответствуют техническим мерам защиты информации, включая технические, криптографические и программно-аппаратные методы и средства [7, 8].

Исследование влияния вероятностных моделей на показатели эффективности.

В таблице 1 приведены значения числовых характеристик для пяти выбранных работ (значения исходных данных выбраны для проведения данного исследования совместно с экспертами).

Таблица 1. Значения числовых характеристик

j	1	2	3	4	5
mt_j , сут.	30,0	60,0	30,0	5,0	45,0
kvt_j	0,20	0,10	0,10	0,10	0,10
mz_j , тыс. руб.	700,0	1000,0	200,0	120,0	500,0
kvz_j	0,25	0,25	1,25	0,30	0,15

В таблице 2 приведены три варианта (В) вероятностных моделей для интервалов и затрат этих работ: Норм. – нормальное распределение; Бета – бета распределение; Равн. – равномерное

ное распределение; Логнорм. – логарифмически нормальное распределение; Б.-С. – распределение Бирнбаума-Саундерса; Парето – распределение Парето с нулевой точкой; Гамма – гамма-распределение. Эти распределения используются при имитационном моделировании, в теории рисков и в страховой математике [5, 6, 9].

Таблица 2. Вероятностные модели (В. м.)

В	j	1	2	3	4	5
1.1	В. м. для интер.	Норм.	Бета	Равн.	Бета	Норм.
	В. м. для затрат	Логнорм.	Б.-С.	Парето	Логнорм.	Гамма
	kvz_j	0,25	0,25	1,25	0,30	0,15
1.2	В. м. для интер.	Бета	Норм.	Бета	Равн.	Норм.
	В. м. для затрат	Б.-С.	Логнорм.	Логнорм.	Парето	Гамма
	kvz_j	0,25	0,25	0,30	1,25	0,15
1.3	В. м. для интер.	Бета	Бета	Норм.	Равн.	Равн.
	В. м. для затрат	Парето	Гамма	Логнорм.	Б.-С.	Б.-С.
	kvz_j	1,25	0,25	0,25	0,30	0,15

Значения интервалов между поступлениями платежей (h_j) и их величин (Y_{0j}) для выбранных пяти работ приведены в таблице 3.

Таблица 3. Значения интервалов между поступлениями платежей и их величин

j	1	2	3	4	5
h_j , сут.	30,0	60,0	10,0	5,0	45,0
Y_{0j} , тыс. руб.	700,0	1000,0	66,67	120,0	500,0

В таблицах 4, 5, 6 приведены результаты вычисления показателей эффективности для исходных данных из таблиц 1, 2 и 3 по трем вариантам. Эти результаты приведены для двух значений коэффициента g (2), влияющего на начальное значение процесса (1).

Таблица 4. Результаты моделирования (В 1.1)

g	S_t , сут.	k_τ	\tilde{R}_τ	τ_1	τ_2	\tilde{c}_v , %	v_1 , %	v_2 , %
0,05	90	5644	0,2822	0,2770	0,2875	23,22	22,17	24,31
	180	10174	0,5087	0,5029	0,5145	48,31	47,16	49,48
	270	11417	0,5708	0,5651	0,5766	55,21	54,04	56,40
	360	12833	0,6417	0,6360	0,6472	65,76	64,54	67,01
0,10	90	183	0,0092	0,0081	0,0103	25,98	20,36	33,16
	180	1294	0,0647	0,0619	0,0676	29,34	26,91	31,98
	270	1904	0,0952	0,0918	0,0987	32,90	30,76	35,19
	360	3160	0,1580	0,1538	0,1623	37,93	36,13	39,82

Таблица 5. Результаты моделирования (В 1.2)

g	S_t , сут.	k_τ	\tilde{R}_τ	τ_1	τ_2	\tilde{c}_v , %	v_1 , %	v_2 , %
0,05	90	6395	0,3197	0,3143	0,3252	28,41	27,31	29,55
	180	10693	0,5346	0,5288	0,5405	51,09	49,94	52,27
	270	11931	0,5966	0,5908	0,6023	58,68	57,50	59,89
	360	13190	0,6595	0,6539	0,6650	69,02	67,78	70,29
0,10	90	411	0,0205	0,0189	0,0223	24,93	21,12	29,42
	180	2072	0,1036	0,1001	0,1072	32,65	30,61	34,83
	270	2935	0,1467	0,1426	0,1509	36,15	34,33	38,06

g	S_t , сут.	k_τ	\tilde{R}_τ	τ_1	τ_2	\tilde{c}_v , %	v_1 , %	v_2 , %
	360	4448	0,2224	0,2176	0,2273	41,88	40,28	43,55

Таблица 6. Результаты моделирования (В 1.3)

g	S_t , сут.	k_τ	\tilde{R}_τ	τ_1	τ_2	\tilde{c}_v , %	v_1 , %	v_2 , %
0,05	90	6503	0,3251	0,3197	0,3306	36,37	35,14	37,65
	180	10169	0,5084	0,5026	0,5143	56,05	54,80	57,32
	270	11541	0,5770	0,5713	0,5828	64,68	63,40	65,98
	360	12662	0,6331	0,6275	0,6387	73,65	72,32	75,00
0,10	90	1916	0,0958	0,0924	0,0993	35,92	33,69	38,30
	180	4370	0,2185	0,2137	0,2234	47,37	45,65	49,16
	270	5742	0,2871	0,2818	0,2924	52,09	50,51	53,73
	360	7069	0,3534	0,3479	0,3591	57,43	55,91	58,98

Из результатов, приведенных в таблицах (4 – 6), видно, что виды вероятностных моделей влияют на значения обоих показателей эффективности. Несовпадения значений статистически значимо, так как отсутствуют пересечения доверительных интервалов. Исключения составили три случая для коэффициента вариации в таблицах 4 и 5, когда $g = 0,10$. В таблице 5 эти случаи выделены жирным шрифтом. Это связано с тем, что при увеличении начального значения процесса (1) уменьшаются объемы выборочных значений k_τ , а это приводит к расширению доверительных интервалов.

Исследование влияния значений коэффициентов вариации на показатели эффективности. В таблице 7 приведены выбранные для исследования вероятностные модели и значения математических значений для интервалов и затрат. Исходные данные для поступлений в БФ приведены в таблице 3.

Таблица 7. Вероятностные модели и значения математических ожиданий

j	1	2	3	4	5
В. м. для интер.	Норм.	Бета	Равн.	Бета	Норм.
mt_j , сут.	30,0	60,0	30,0	5,0	45,0
В. м. для затрат	Логнорм.	Б.-С.	Парето	Логнорм.	Гамма
mz_j , тыс. руб.	700,0	1000,0	200,0	120,0	500,0

В таблице 8 приведены три варианта (В) значений коэффициентов вариации для интервалов и затрат для данного исследования: 1-й вариант – это значения из таблицы 1; 2-й вариант – когда коэффициенты вариации меньше, чем в первом; 3-й вариант – когда коэффициенты вариации больше, чем в первом.

Таблица 8. Вероятностные модели (В. м.)

В	j	1	2	3	4	5
2.1	kvt_j	0,20	0,10	0,10	0,10	0,10
	kvz_j	0,25	0,25	1,25	0,30	0,15
2.2	kvt_j	0,10	0,05	0,05	0,05	0,05
	kvz_j	0,15	0,15	1,10	0,20	0,05
2.3	kvt_j	0,30	0,15	0,15	0,15	0,15
	kvz_j	0,35	0,35	1,40	0,40	0,25

В таблицах 9, 10 приведены результаты вычисления показателей эффективности для исходных данных из таблиц 3, 7 и 8 по двум вариантам соответственно. Эти результаты также

приведены для двух значений коэффициента g (2), что повышает качество выводов. Для варианта 2.1 результаты приведены в таблице 4.

Таблица 9. Результаты моделирования (В 2.2)

g	S_t , сут.	k_r	\tilde{R}_r	τ_1	τ_2	\tilde{c}_v , %	v_1 , %	v_2 , %
0,05	90	5341	0,2671	0,2619	0,2722	19,86	18,87	20,89
	180	9778	0,4889	0,4831	0,4947	47,99	46,82	49,18
	270	10943	0,5472	0,5413	0,5530	54,18	53,00	55,38
	360	12446	0,6223	0,6166	0,6280	64,64	63,41	65,89
0,10	90	64	0,0032	0,0026	0,0039	22,17	14,19	34,65
	180	634	0,0317	0,0297	0,0338	24,29	21,22	27,80
	270	909	0,0454	0,0431	0,0479	28,03	25,23	31,14
	360	1737	0,0868	0,0836	0,0902	34,61	32,32	37,07

Таблица 10. Результаты моделирования (В 2.3)

g	S_t , сут.	k_r	\tilde{R}_r	τ_1	τ_2	\tilde{c}_v , %	v_1 , %	v_2 , %
0,05	90	6566	0,3283	0,3228	0,3338	27,59	26,53	28,70
	180	10873	0,5436	0,5378	0,5495	50,91	49,77	52,08
	270	12080	0,6040	0,5983	0,6097	58,83	57,65	60,03
	360	13337	0,6669	0,6613	0,6723	69,44	68,19	70,70
0,10	90	493	0,0246	0,0229	0,0265	25,73	22,17	29,87
	180	2408	0,1204	0,1166	0,1243	33,02	31,11	35,05
	270	3442	0,1721	0,1677	0,1766	37,55	35,83	39,35
	360	4935	0,2467	0,2417	0,2518	43,50	41,95	45,11

Из результатов, приведенных в таблицах 4, 9 и 10, видно, что значения коэффициентов вариации влияют на значения обоих показателей эффективности, но на значения коэффициента вариации (8) не всегда значимо. Подчеркнем, что не надо путать коэффициент вариации (8), как показатель эффективности, и коэффициенты вариации для исходных данных, влияние которых мы исследуем. Так, например, при увеличении значений коэффициентов вариации для исходных данных показатель (8) для одного случая изменился не значимо. В таблице 10 этот случай выделен жирным шрифтом, сравнение делается с результатами из таблицы 4. При уменьшении этих коэффициентов таких случаев стало больше, в таблице 9 эти случаи выделены жирным шрифтом. При сравнении результатов из таблиц 9 и 10 все показатели эффективности (8) статистически различимы, кроме случая, когда $g = 0,10$ и $S_t = 90$. Это связано с тем, что для этого случая малый объем выборки, $k_r = 64$ (таблица 9), что увеличивает доверительный интервал.

Заключение. С использованием предложенного программно-математического обеспечения на основе дискретно-имитационного моделирования показано, что виды вероятностных моделей и их числовые характеристики, описывающие интервалы времени между инцидентами информационной безопасности и затраты на их устранение, оказывают влияние на значения показателей эффективности в виде вероятности «обнуления» бюджетного фонда и коэффициента вариации. В связи с этим в моделирующей программе необходимо расширять список вероятностных моделей и улучшать интерфейс для ее настройки на реальный случай. Снижение значений показателей эффективности до допустимых значений можно осуществлять за счет увеличения начального значения F_{s_0} процесса (1). С увеличением этого значения

показатели эффективности уменьшаются. Выбор этого значения определяется по результатам имитационного моделирования.

Список источников

1. Оганесян Л.Л. Проектное управление в информационной безопасности / Л.Л. Оганесян, Н.С. Козырь // Вестник Академии знаний, 2023. – № 4(57). – С. 207-209.
2. Сизов В.А. Моделирование экономики информационной безопасности субъекта экономической деятельности на основе симплекс-метода / В.А. Сизов, А.А. Дрожкин // Вестник Российского экономического университета имени Г.В. Плеханова, 2021. – Т. 18. – № 1(115). – С. 173-178.
3. Ефимов Е.Н. Оценка эффективности мероприятий информационной безопасности в условиях неопределенности / Е.Н. Ефимов, Е.М. Лапицкая // Бизнес-информатика, 2015. – № 1(31). – С. 51-57.
4. Краковский Ю.М. Программно-математическое обеспечение для исследования показателей эффективности экономики информационной безопасности / Ю.М. Краковский, В.П. Киргизбаев // System analysis and mathematical modeling. – Иркутск, Байкальский государственный университет, 2024. – Т. 6. – № 2. – С. 209–220. – DOI 10.17150/2713-1734.2024.6(2).209-220
5. Кельтон В. Имитационное моделирование / В. Кельтон, А. Лоу. – СПб.: Питер, 2004. – 847 с.
6. Краковский Ю.М. Моделирование ремонтных работ оборудования на основе случайного процесса риска / Ю.М. Краковский, Н.А. Хоанг // Прикладная информатика, 2020. – Т.15. – № 6. – С. 5–15. DOI:10.37791/2687-0649-2020-15-6-5-15.
7. Нестеров С.А. Основы информационной безопасности / С.А. Нестеров. — Санкт-Петербург: Лань, 2023. – 324 с.
8. Краковский Ю.М. Методы защиты информации / Ю.М. Краковский. – Санкт-Петербург: Лань, 2021. – 236 с.
9. Мак Томас. Математика рискованного страхования / Томас М. Пер. с нем. – М.: ЗАО «Олимп-Бизнес», 2005. – 432 с.

Краковский Юрий Мечеславович. Доктор технических наук, профессор кафедры «Информационные системы и защита информации», Иркутский государственный университет путей сообщения, 79149267772@yandex.ru, 664074, Россия, г. Иркутск, ул. Чернышевского, д. 15.

Киргизбаев Владислав Павлович. Аспирант кафедры «Информационные системы и защита информации», Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация, v.p.kirgizbaev@gmail.com, 664074, Россия, г. Иркутск, ул. Чернышевского, д. 15.

UDC 519.237.5

DOI:10.25729/ESI.2024.35.3.010

The impact of probabilistic models of works related to information protection on the values of performance indicators

Yuri M. Krakovsky, Vladislav P. Kirgizbaev

Irkutsk state transport university, Russia, Irkutsk, 79149267772@yandex.ru

Abstract. To enhance the efficiency of information protection tasks within an organization, it is proposed to use a budgetary fund where necessary financial resources are accumulated and then expended. Under conditions of uncertainty, both the time intervals for expenditures and the expenditures themselves are random variables, which should be described by probabilistic models when modeling. A study was conducted on the impact of types of probabilistic models and the values of their numerical characteristics on the efficiency indicators of tasks performed by the organization's information security service staff, servicing the corporate information system. The efficiency indicators are the probability of "zeroing" the budgetary fund and the coefficient of variation, which are replaced by point and interval estimates in discrete-event simulation. Practical recommendations have been obtained.

Keywords: discrete-event simulation, budgetary fund, economics of information security, efficiency indicators

References

1. Oganesyanyan L.L., Kozyr N.S. *Proyektnoye upravleniye v informatsionnoy bezopasnosti* [Project management in information security]. *Vestnik Akademii znaniy* [Bulletin of the Academy of Knowledge], 2023, no. 4(57), pp. 207-209.
2. Efimov E.N., Lapitskaya E.M. *Modelirovaniye ekonomiki informatsionnoy bezopasnosti sub"yekta ekonomicheskoy deyatelnosti na osnove simpleks-metoda* [Assessment of the effectiveness of information security measures in conditions of uncertainty]. *Biznes-informatika* [Business Informatics], 2015, no. 1(31), pp. 51-57.
3. Sizov V.A., Drozhkin A.A. *Otsenka effektivnosti meropriyatiy informatsionnoy bezopasnosti v usloviyakh neopredelennosti* [Modeling the economics of information security for an economic entity based on the simplex method]. *Bulletin of the Russian Economic University named after G.V. Plekhanov*, 2021, vol. 18, no. 1(115), pp. 173-178.
4. Krakowski Y.M., Kirgizbaev V.P. *Programmno-matematicheskoye obespecheniye dlya issledovaniya pokazateley effektivnosti ekonomiki informatsionnoy bezopasnosti* [Software and mathematical tools for investigating the efficiency indicators of information security economics]. *System analysis and mathematical modeling*, Irkutsk, Baikal state university, 2024, vol. 6, no. 2, pp. 209–220, DOI:10.17150/2713-1734.2024.6(2).209-220
5. Kelton, W., Law A. *Imitatsionnoye modelirovaniye* [Simulation modeling]. Saint Petersburg, Peter, 2004, 847 p.
6. Krakovsky Yu.M., Hoang N.A. *Modelirovaniye remontnykh rabot oborudovaniya na osnove sluchaynogo protsessa riska* [Modeling equipment repair works based on a random risk process]. *Prikladnaya informatika* [Applied informatics], 2020, vol. 15, no. 6. pp. 5–15, DOI:10.37791/2687-0649-2020-15-6-5-15.
7. Nesterov S.A. *Osnovy informatsionnoy bezopasnosti* [Fundamentals of information security]. St. Petersburg, Lan, 2023, 324 p.
8. Krakovsky Yu.M. *Metody zashchity informatsii* [Information security methods]. St. Petersburg, Lan, 2021, 236 p.
9. Thomas Mack. *Matematika riskovogo strakhovaniya* [Mathematics of risk insurance]. Moscow, Olimp-Business, 2005, 432 p.

Krakovsky Yuri Mecheslavovich. *Doctor of technical sciences, professor of the department of information systems and information security, Irkutsk state transport university, 79149267772@yandex.ru, 664074, Russia, Irkutsk, Chernyshevskogo str., 15.*

Kirgizbaev Vladislav Pavlovich. *PhD student of the department of information systems and security information, Irkutsk state transport university, v.p.kirgizbaev@gmail.com, 664074, Russia, Irkutsk, Chernyshevskogo str., 15.*

Статья поступила в редакцию 11.03.2024; одобрена после рецензирования 26.08.2024; принята к публикации 08.10.2024.

The article was submitted 03/11/2024; approved after reviewing 08/26/2024; accepted for publication 10/08/2024.

УДК 004.896

DOI:10.25729/ESI.2024.35.3.011

Формирование базы правил для PLM систем

Камалетдинова Лилия Рашидовна, Романов Антон Алексеевич, Филиппов Алексей Александрович, Ярушкина Надежда Глебовна

Ульяновский государственный технический университет,
Россия, Ульяновск, *lribragimova1995@gmail.com*

Аннотация. Часто возникает необходимость интеграции PLM-системы (Product Lifecycle Management) и сторонней информационной системы, не входящей в состав комплекса PLM, но участвующей в решении задач, связанных с информационной поддержкой управления производственными процессами. На данный момент задачу такой интеграции выполняет аналитик. Он должен сформировать структурную и процессную модели интегрируемой ИС, на основе которых формируются правила взаимодействия с системой. В процессе эксплуатации, помимо аналитика, участвует оператор (выполняет задачу поддержания актуальности данных, как интегрируемой ИС, так и всего комплекса) и лицо, принимающее решение (ЛПР) (выполняет процесс управления интегрируемой ИС, а именно поддержание её в работоспособном состоянии). Предлагается подход, позволяющий снизить нагрузку как на аналитика, обеспечивающего настройку взаимодействия и отображения данных, так и на оператора, участвующего в обеспечении актуальности структур данных, и ЛПР, участвующего в принятии важных решений, связанных с возможными рисками на производстве.

Ключевые слова: база правил, PLM-системы, управление на основе данных

Цитирование: Камалетдинова Л.Р. Формирование базы правил для PLM систем / Л.Р. Камалетдинова, А.А. Романов, А.А. Филиппов, Н.Г. Ярушкина // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 120-132. – DOI: 10.25729/ESI.2024.35.3.011.

Введение. На сегодняшний день многие крупные производства используют PLM системы [1] (Product Lifecycle Management) для управления жизненным циклом изделий. PLM системы используются для осуществления контроля потоков данных на производстве, включая решение таких задач, как хранение, интеграцию, поддержание в актуальном состоянии данных в каждой информационной подсистеме, входящей в комплекс. Ключевую роль в этом выполняет шина данных [2]. В каждом комплексе систем шина данных имеет свои особенности реализации. Однако общей особенностью интеграции на основе шин данных является использование взаимодействия, основанного на правилах [3]. Для настройки взаимодействия между подсистемами аналитик выполняет:

- построение модели поведения информационной системы,
- выявление ключевых признаков,
- формирование правил поведения системы.

В статье рассматривается не только подход к формированию взаимодействия между системами, но и подход к управлению PLM подсистемами. Часто возникает необходимость интеграции PLM-системы и сторонней информационной системы (далее в тексте будем называть ее интегрируемая ИС), не входящей в состав комплекса PLM, но участвующей в решении задач, связанных с информационной поддержкой управления производственными процессами. На данный момент задачу такой интеграции выполняет аналитик или лицо принимающее решение (ЛПР). Аналитик должен сформировать структурную и процессную модели интегрируемой ИС, на основе которых формируются правила взаимодействия с системой. Также в процессе эксплуатации к задачам аналитика (ЛПР) добавляется процесс управления интегрируемой ИС, а именно поддержание актуальности данных и работоспособного состояния самой системы.

На рисунке 1 представлен граф состояний интегрируемой ИС в общем виде без участия человека. На рисунке представлены основные состояния интегрируемой ИС, используемые в

данной работе. Так, $S1$ — работоспособное состояние системы, $S2$ — неработоспособное состояние системы или состояние отказа, $S3$ — предельное состояние системы [4].

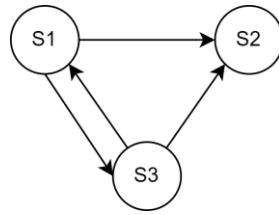


Рис. 1. Граф состояний интегрируемой информационной системы

Под работоспособным состоянием системы понимается такое состояние, когда система продолжает выполнять свои основные задачи, при этом данные, используемые системой, соответствуют нормальному состоянию данных для этой системы. Под предельным состоянием системы понимается такое состояние системы, когда система продолжает выполнять свои основные задачи, при этом данные системы находятся в пограничном состоянии, то есть на границе нормы. Под неработоспособным состоянием системы или состоянием отказа понимается такое состояние системы, когда система не может продолжать выполнение основных задач и требуется вмешательство ЛПР. Находясь в состоянии $S1$ и $S3$, система может перейти в состояние отказа, поэтому основная цель управления состоит в поддержании системы в состоянии $S1$ и возможности возврата в это состояние без рисков, в случае, если система перешла в состояние $S2$, вернуться в рабочий цикл ($S1 - S3$) возможно лишь при помощи ЛПР.

Стоит отдельно рассмотреть процесс, когда взаимодействие интегрируемой ИС и всего комплекса PLM обеспечивается участием оператора. В этом процессе в настоящее время объем работ оператора для настройки взаимодействия, отображения данных, поддержания системы в работоспособном состоянии достаточно велик. Роль оператора в данном процессе отображена на рисунке 2.

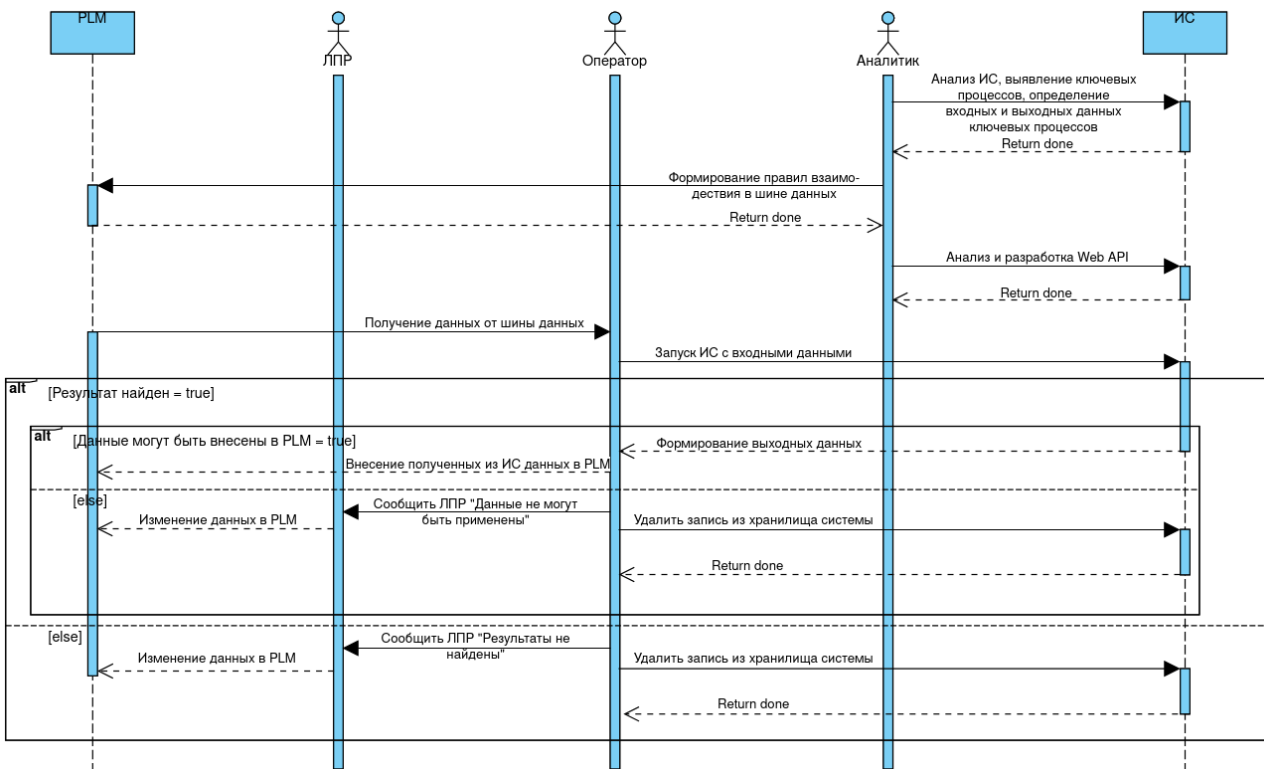


Рис. 2. Процесс организации взаимодействия и поддержания системы в работоспособном состоянии с участием человека

Аналитик должен учитывать изменения структуры данных интегрируемой ИС. Оператор отслеживает возможные риски, связанные с обеспечением соответствия данным комплекса PLM. ЛПП принимает решения относительно возможной минимизации рисков. Предлагается подход, позволяющий снизить нагрузку как на аналитика, обеспечивающего настройку взаимодействия и отображения данных, так и на оператора, участвующего в обеспечении актуальности структур данных, и ЛПП, участвующего в принятии важных решений, связанных с возможными рисками на производстве.

Обзор существующих методов и подходов к управлению. В настоящее время подходы к управлению производственными ИС раскрываются в публикациях [5-10]. В источниках [5-6, 8] предлагается использовать модель интегрируемой ИС. Данные подходы напрямую зависят от точности и полноты модели, однако формирование модели реализуется либо аналитиком [5, 8], либо автоматически [6], что не всегда гарантирует адекватность и точность модели. В [5] представлено классическое управление на основе модели процессов и данных информационной системы, где роль аналитика велика. В [6] предлагается использовать фреймворк для формирования модели. Несмотря на то, что процесс формирования модели происходит автоматически, управляющий компонент всё ещё зависит от точности и полноты модели. В [7] предлагается использовать обучающую выборку и машинное обучение без учителя. Однако, формируемая модель интегрируемой ИС зависит от полноты и качества данных выборки, что зачастую может привести к неверному истолкованию результатов прогнозирования механизмов управления, в случае, если данные имеют ошибки или сильно зашумлены. В [8] предлагается подход, основанный на использовании цифровых двойников. Несмотря на преимущества данного подхода в целом, применительно к поставленной задаче (интеграция информационных систем, отображение данных и снижение трудозатрат аналитика, оператора и ЛПП) подход, основанный на использовании цифровых двойников, требует дополнительного качественного анализа и построения модели информационной системы, что приводит к увеличению нагрузки на аналитика.

В [9] описан подход к управлению human-in-the-loop. Основная идея данного подхода заключается в полном контроле поведения системы управления со стороны человека и сводится к машинному обучению с учителем. Данный подход является наиболее безопасным с точки зрения предупреждения рисков, связанных с производством, но более комплексным и требующим участия ЛПП на всех этапах жизненного цикла управляющей системы, начиная с формирования обучающей выборки, и заканчивая контролем поведения управляющей системы.

Таким образом, предлагается использовать подход управления на основе данных [10]. Предполагается снижение трудозатрат аналитика и оператора за счет выполнения задачи интеграции и частично задачи управления интегрируемой ИС разработанной управляющей системой. Подход управления на основе данных подразумевает:

- моделирование системы, но не в классическом понимании [5], а за счет формирования метамоделей данных интегрируемой ИС на основе анализа ее хранилища;
- отображение данных при интеграции систем PLM и интегрируемой ИС предприятия за счет использования базы правил поведения интегрируемой ИС;
- снижение рисков перехода в состояние отказа (рисунок 1) и поддержание системы в работоспособном состоянии без участия ЛПП в данном процессе.

Обзор существующих подходов к формированию баз правил. Для решения задачи предлагается использовать продукционную модель базы знаний. Классические нечеткие системы основаны на подходе Мамдани [11]. В таких системах существуют 2 модуля преобразования обычных данных в нечеткие. Модуль фаззификации устанавливает соответствие

между реальными значениями входных данных и нечеткими значениями, основываясь на функции принадлежности. С другой стороны, модуль дефаззификации устанавливает соответствие между нечеткими значениями и реальными значениями выходных данных предметной области. Нечеткие продукционные базы знаний (системы, основанные на нечетких правилах) основаны на принципе преобразования четких значений в нечеткие. Правила в данном случае представляют собой множество лингвистических термов и выходные данные ассоциируются с ними, например, правило может иметь несколько входных данных и только одно определенное значение выходного параметра. Существует несколько вариантов классической системы, основанной на нечетких правилах:

1. Система, основанная на нечетких правилах Мамдани [12]. Каждая переменная правила представляет собой значение из множества лингвистических термов, например, если X_i – переменная, представленная множеством $\{l_1, l_2, l_3\}$, тогда в переменной правила X_i может быть представлена как $\{l_1, l_2\}$. Переменная может принадлежать набору лингвистических терминов в правиле. Это помогает сократить количество правил, чтобы избежать проблемы увеличения размера самой базы. Таким образом, правило может иметь вид:

$$x_1 = \{l_{11}, l_{12}\} \wedge x_2 = \{l_{23}\} \wedge \dots \wedge x_n = \{l_{n1}, l_{n2}\} \rightarrow y = Y.$$

2. Приближенные системы, основанные на нечетких правилах Мамдани [13]. Такие системы включают в себя несколько элементов наборов терминов, которые могут снизить интерпретируемость вывода. Системы способны добиться большей точности ценой потери интерпретируемости. Каждое правило имеет свой собственный нечеткий набор вместо использования лингвистических терминов. Этот подход генерирует семантически свободные правила и обладает более высокой выразительностью благодаря использованию различных нечетких множеств в каждом правиле. Он может принимать разное количество правил в зависимости от сложности проблемы. Что касается недостатков, подобные системы страдают от потери интерпретируемости, а также они могут переопределять обучающие данные и плохо работать в случае скрытых зависимостей признаков в данных.
3. Системы классификации на основе нечетких правил [14]. Система классификации на основе нечетких правил – это система, которая использует нечеткие правила в качестве средства обучения. В классических системах, основанных на нечетких правилах Мамдани, входные данные сопоставляются обычно с одномерным выходом, но в данном случае входные данные сопоставляются с одной из меток класса. Структура правила выглядит следующим образом:

$$x_1 = l_{1i} \wedge x_2 = l_{2i} \wedge \dots \wedge x_n = l_{ni} \rightarrow y = c.$$

Существуют варианты неклассических систем, основанных на нечетких правилах:

1. Иерархические нечеткие системы [15] состоят из нескольких низкоразмерных нечетких систем, расположенных иерархическим образом. Правила в иерархических нечетких системах сгруппированы в модули в соответствии с их ролями в системе. Каждый модуль вычисляет частичное решение, которое далее передается на модули следующего уровня. Хотя каждый модуль представляет собой нечеткую систему, он генерирует значительно меньшее количество правил, чем плоская нечеткая система. Несмотря на широкое распространение применения иерархических баз правил при работе с big data, существует ряд недостатков подобных систем. В [16] приведены примеры оптимизации иерархических нечетких систем с помощью генетического алгоритма. В [17] приведены примеры оптимизации каскадных иерархических нечетких систем с использованием нейронных сетей, однако, из-за того, что предложенный метод задействует все входные переменные, теряется преимущество сокращения количества правил.
2. Нейро-нечеткие системы [18] представляют собой слияние систем, основанных на нечетких правилах, с искусственными нейронными сетями. Основная идея систем состоит в способности принятия решений на основе заданных правил и обучения за счет использования нейронных систем. База правил является плоской (одномерной, в отличие от

иерархических нечетких систем), что увеличивает размеры самой базы правил, но упрощает процесс обучения нейронной сети в процессе работы системы управления. Нейро-нечеткие системы состоят из двух модулей, первый отвечает за настройку и структурирование условия правила, а второй отвечает за формирование следствия. Нейронная сеть на основе существующих правил, принимая входные данные, настраивает условие правила и частично вычисляет функцию принадлежности, соответствующую нечеткому множеству входных данных. На втором этапе происходит вычисление следствия правила от нечеткого множества следствий до точных значений вывода.

3. Эволюционные нечеткие системы [19]. Для решения задачи формирования базы правил используются генетические алгоритмы. Отличие данного типа нечетких баз правил состоит в том, что она самообучаема и самооптимизируема. Реализация этого подхода выполняется в два этапа:

- на первом этапе генетический алгоритм (ГА) используется для нахождения кандидатов нечетких правил в базе знаний;
- на втором этапе с помощью ГА выполняется процедура оптимизации базы знаний с целью исключения из рассмотрения худших правил, мало влияющих на получение окончательного решения.

Для решения задачи исследования данных в организации было принято решение объединить подходы к построению иерархической нечеткой базы правил и нейро-нечетких баз правил с выводом результата, основанном на нечетких правилах Мамдани, в рамках подхода, использующего эволюционный алгоритм для формирования правил на основе изменений метамодели. Таким образом, база правил будет построена, как показано на рисунке 3.

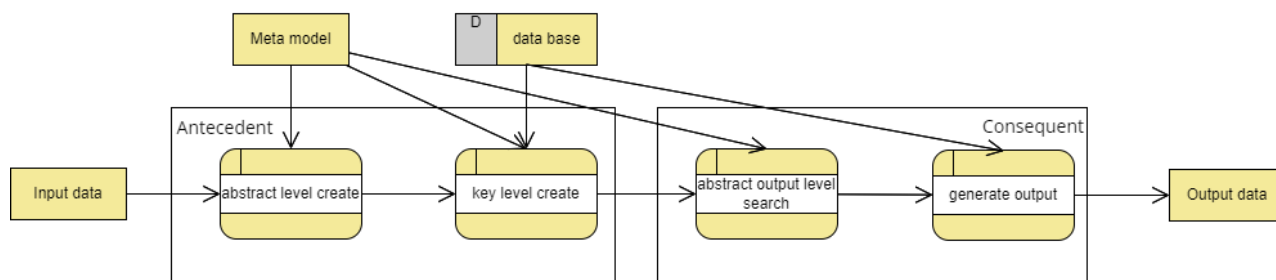


Рис. 3. Структура гибридной базы знаний с нечетким логическим выводом

Структура гибридной базы знаний с нечетким логическим выводом, представленной на рисунке 3, состоит из 2-х условных блоков.

Первый блок – Antecedent – блок формирования условий правил, который состоит из построения абстрактного уровня правил (abstract level create) и построения ключевого уровня (key level create), содержащего условия базы правил второго уровня. При построении абстрактного уровня правил участвует сформированная метамодель интегрируемой информационной системы (meta model), а при построении ключевого уровня условий базы правил второго уровня участвует, помимо метамодели, база данных интегрируемой информационной системы (data base).

Второй блок – Consequent – блок построения следствий правил, который состоит из формирования абстрактного вывода базы правил второго уровня (abstract output level search) и генерации нечеткого логического вывода (generate output) на основе нечетких термов, полученных из предыдущего процесса. При формировании абстрактного вывода задействована метамодель интегрируемой информационной системы (meta model), а при генерации четкого логического вывода участвует только база данных интегрируемой информационной системы (data base).

Так, процесс работы базы правил получает на входе данные (input data) от шины данных PLM системы, а на выходе данные (output data) в формате, подходящем для интеграции с базой данных интегрируемой информационной системы.

Таким образом, согласно диаграмме, база знаний состоит из нескольких уровней условий, учитывающих возможные изменения в метамодели и в базе данных интегрируемой информационной системы, и нескольких уровней следствий, формирующих четкий логический вывод на основе лингвистического представления правил в базе правил.

Модель базы знаний поведения интегрируемой информационной системы. Согласно рисунку 3, база правил будет представлена в виде иерархической структуры с двумя уровнями правил. Ранее авторами была получена структурная модель метаданных M интегрируемой ИС [20]. Так, первый уровень будет представлен в виде правил, состоящих из лингвистических термов, и зависеть от изменений метамодели. Второй уровень правил будет формироваться динамически на основе результатов, полученных на первом уровне.

Пусть $INP = \{INP_1, INP_2, \dots, INP_z\}, z \in N$ – множество лингвистических термов, представляющих входные данные модели метаданных M , а

$OUT = \{OUT_1, OUT_2, \dots, OUT_w\}, w \in N$ – множество лингвистических термов, представляющих ключевые процессы модели метаданных M , тогда правило, описывающее первый уровень, будет иметь теоретико-множественное представление в виде:

$$P(INP) \rightarrow \{INP^{OUT_s}\}, OUT_s, \tag{1}$$

где OUT_s – лингвистический терм, отражающий конкретный ключевой процесс модели метаданных M , а $\{INP^{OUT_s}\}$ – множество лингвистических термов, отражающих входные данные для конкретного ключевого процесса модели метаданных M .

Представим $\{INP^{OUT_s}\}$ как X , а OUT_s как Y . В таблице 1 представлены входные данные для правила второго уровня, которые будут использованы для формирования конечных значений поведения интегрируемой ИС.

Таблица 1. Представление входных данных для правила второго уровня

x_1	x_2	...	x_m	y
v^1_1	v^1_2	...	v^1_m	y^1
v^2_1	v^2_2	...	v^2_m	y^2
...
v^n_1	v^n_2	...	v^n_m	y^n

Колонки 1-4 отображают значения входных данных для ключевых процессов информационной системы, а колонка 5 представляет значения данных, которые являются реакцией системы на значения входных данных.

Для формирования правил второго уровня управления интегрируемой ИС определим следующий функционал:

$$p(X, V) \rightarrow Y, \tag{2}$$

где $V = \{\{v^1_1, \dots, v^1_m\}, \{v^2_1, \dots, v^2_m\}, \dots, \{v^n_1, \dots, v^n_m\}\}$ – входные значения m параметров,

$$X = \{x_k\}, k = [1, m], m \in N, n \in N,$$

$$Y = \{y^i\}, i = [1, n], n \in N, - n \text{ состояний интегрируемой ИС.}$$

Состояние системы y^i определяется вектором входных значений $\{v^i_1, \dots, v^i_m\}$. Таким образом, для формирования правила управления системой для перевода ее в состояние “выдача управляющих воздействий” y^i необходимо в antecedent правила включить сравнение вектора параметров X со значениями $\{v^i_1, \dots, v^i_m\}$:

$$p^i(X, \{v^i_1, \dots, v^i_m\}) \rightarrow y^i. \tag{3}$$

При работе интегрируемой ИС могут возникать ситуации, когда разные значения входных параметров приводят к одному и тому же состоянию y^i . В этом случае их необходимо группировать, явно указав одно и то же выходное состояние.

Для учета неопределенности во входных значениях будем использовать нечеткие функции принадлежности треугольной формы $\mu_{y,i}(x^i)$ [11]. Данная функция значений входных параметров x^i , присущих состоянию системы i , позволяет выполнять логический вывод даже в том случае, когда вектор входных значений содержит значения, не совпадающие в точности со значениями, используемыми в антецедентах правил.

Алгоритм формирования выходных данных на основе иерархической базы правил.

На рисунке 4 представлен алгоритм принятия решения с использованием иерархической нечеткой базы правил с нечетким логическим выводом, основанном на подходе Мамдани [11].

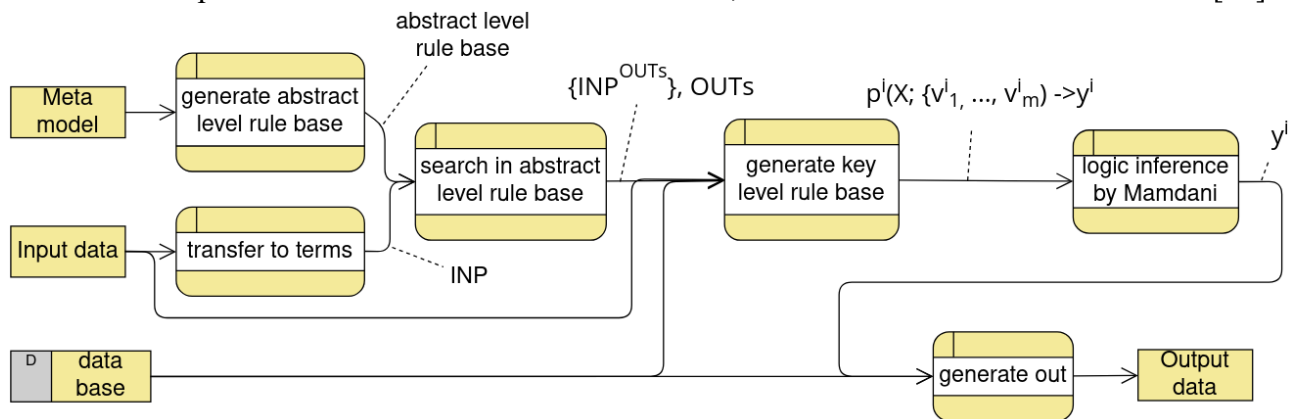


Рис. 4. Алгоритм принятия решений

Предварительно на основе метамодели формируется первый уровень базы правил (generate abstract level rule base), который не будет изменяться до тех пор, пока не произойдут изменения в самой метамодели интегрируемой ИС.

Алгоритм, представленный на рисунке 4, состоит из следующих шагов:

- Входные данные (input data), представленные в виде кортежа данных “ключ-значение” ($inpI = 7$) разного типа (целочисленные, строковые, дата и логические переменные типа boolean), преобразуются в лингвистические термы (transfer to terms), представленные в виде: $INP = \{INP_1, INP_2, \dots, INP_z\}, z \in N$.
- Используя базу правил первого уровня (abstract level rule base) и преобразованные входные данные (INP), осуществляется логический вывод (search in abstract level rule base), представленный в виде: $\{\{INP^{OUT_s}\}, OUT_s\}, s \in N$.
- Результат выполнения правила первого уровня ($\{INP^{OUT_s}\}, OUT_s$), исходные входные данные (input data), представленные в виде кортежа данных “ключ-значение” ($inpI = 7$) разного типа (целочисленные, строковые, дата и логические переменные типа boolean), и база данных интегрируемой ИС (data base) участвуют в динамическом формировании правил второго уровня (generate key level rule base), математическое описание которых представлено в формуле 3.
- В процессе нечеткого логического вывода (logic inference by Mamdani), основанного на подходе Мамдани, получаем результат выполнения правила (y^i) на основе базы правил второго уровня.
- На заключительном этапе (generate out) формируются подходящие выходные данные (output data), представленные в виде “ключ-значение” ($outI = 7$) разного типа (целочисленные, строковые, дата и логические переменные типа boolean). Заключительный этап

использует в качестве входных данных базу данных интегрируемой ИС (data base) и результат выполнения правила (y^i) на основе базы правил второго уровня.

Таким образом, происходит процесс принятия решений на основе иерархической нечеткой базы правил с нечетким логическим выводом.

Иллюстративный пример иерархической базы правил. В качестве примера рассмотрим полученную базу правил на основе информационной системы, предназначенной для обработки заявок, faveo service desk [21]. В работе [22] отображена метамодель данных, представленная на рисунке 5.

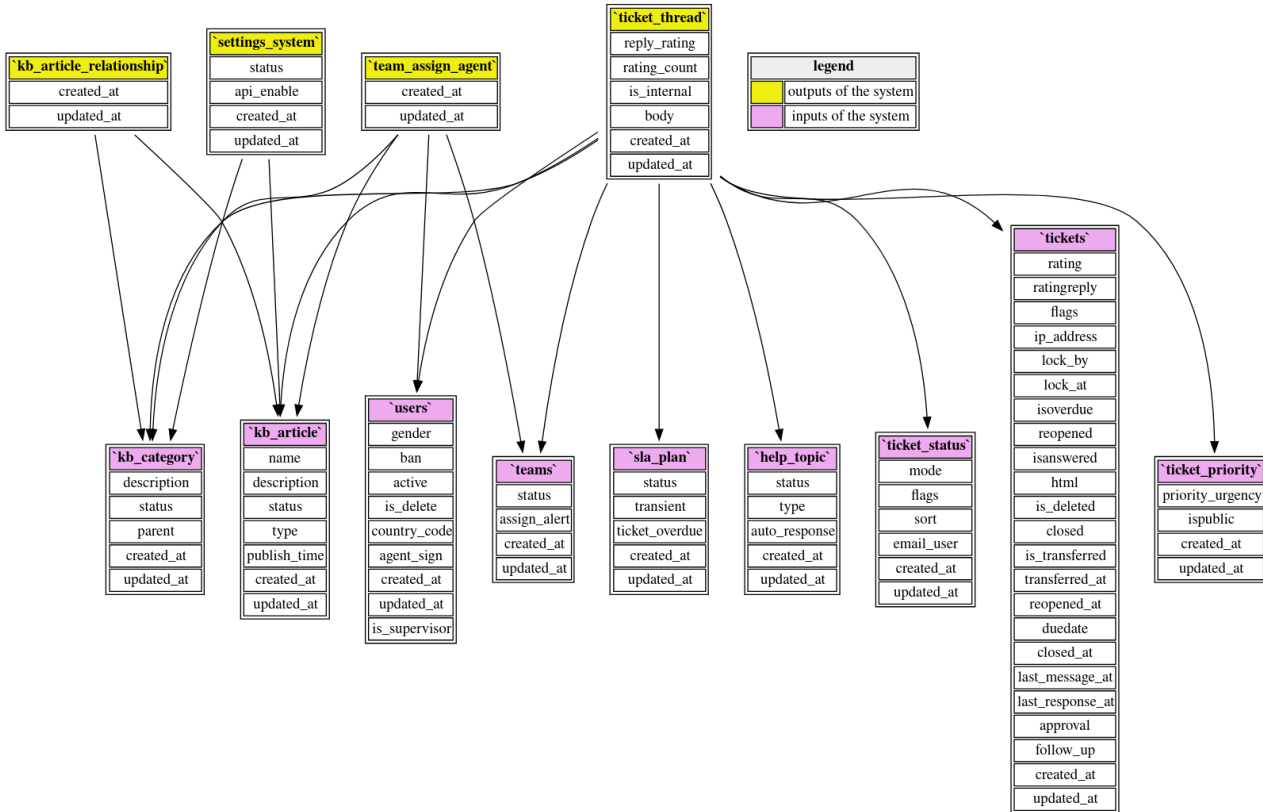


Рис. 5. Метамодель информационной системы faveo

На основе метамодели строится база правил первого уровня, показанная на рисунке 6. Первый уровень содержит лингвистические термы, отображающие зависимость ключевых процессов от входных данных. Входные данные предварительно были преобразованы в лингвистическое описание.

```

IF
INP = {'kb_article', 'kb_category'}
THEN
OUT = {'kb_article', 'kb_category'}, 'kb_article_relationship';
IF
INP = {'timezone', 'time_format', 'date_format', 'date_time_format'}
THEN
OUT = {'timezone', 'time_format', 'date_format', 'date_time_format'}, 'settings_system';
IF
INP = {'teams', 'users'}
THEN
OUT = {'teams', 'users'}, 'team_assign_agent';
IF
INP = {'tickets', 'users', 'ticket_source'}
THEN
OUT = {'tickets', 'users', 'ticket_source'}, 'ticket_thread';
    
```

Рис. 6. Автоматический вывод для отображения базы правил первого уровня информационной системы faveo, полученный на основе метамодели

При сравнении рисунков 5 и 6 видны отличия между названиями входных сущностей для метамодели и входных данных, представленных в виде лингвистических термов в условии

правила. Это связано с тем, что входные данные преобразовываются перед тем, как формировать базу правил первого уровня.

При переходе по ссылке одного из правил, представленных на рисунке 6, динамически формируется база правил второго уровня. На рисунке 7 представлен пример базы правил второго уровня для ключевого процесса `ticket_thread`, отображенного на рисунке 5.

```

IF
ticket_id=14 AND user_id=1 AND source=NULL
THEN id=23;
IF
ticket_id=15 AND user_id=1 AND source=NULL
THEN id=24;
IF
ticket_id=8 AND user_id=1 AND source=NULL
THEN id=32;
IF
ticket_id=15 AND user_id=1 AND source=NULL
ticket_id=14 AND user_id=1 AND source=NULL
ticket_id=13 AND user_id=1 AND source=NULL
ticket_id=12 AND user_id=1 AND source=NULL
ticket_id=11 AND user_id=1 AND source=NULL
ticket_id=10 AND user_id=1 AND source=NULL
ticket_id=9 AND user_id=1 AND source=NULL
THEN id=31;
    
```

Рис. 7. Автоматический вывод для отображения базы правил второго уровня информационной системы `faveo` для ключевого процесса `ticket_thread`

Условия правил содержат ключевые параметры с целочисленными значениями. В процессе нечеткого логического вывода осуществляется поиск подходящего правила с использованием функции принадлежности, в результате которого формируются понятные для пользователя данные, соответствующие состоянию ключевого процесса при переданных входных данных.

Пример вывода результата представлен на рисунке 8.

id	ticket_id	user_id	poster	source	reply_rating	rating_count	is_internal	title	body	format	ip_address	created_at	updated_at
31	9	1			0	0	1		Ticket have been Deleted by Lily Kamaletdinova			2023-11-29 14:27:05	2023-11-29 14:27:05

Рис. 8. Автоматический вывод для отображения результата работы правила 31 рисунка 7

Таким образом, при заданных входных данных формируется результат, который отражает состояние ключевого процесса информационной системы `faveo service desk`. Этот вывод подтверждает целесообразность применения предложенного подхода, так как сформированный результат отражает реальные данные при работе с системой заявок `faveo service desk`, и получение результата не зависит от запуска самой интегрируемой информационной системы, что является преимуществом при отображении данных и параллельной работе интегрируемой информационной системы с комплексом PLM.

Управляющая система в общем комплексе управления. Предлагаемый подход подразумевает использование управляющей системы для сокращения временных затрат оператора, участвующего в процессе управления. Управляющая система (УС) – это программный инструмент, реализующий предлагаемый подход к управлению на основе данных [20]. Управление включает в себя конфигурирование и обмен данными, в данном случае, за счет использования метамодели интегрируемой ИС и иерархической базы правил с нечетким логическим выводом. Управляющая система выступает адаптером для обеспечения взаимодействия интегрируемой ИС с программным комплексом PLM.

На рисунке 9 представлен процесс организации управления информационной системой с участием управляющей системы.

При сравнении рисунков 2 и 9 отметим, что из процесса управления исключаются роли аналитика и оператора, их функции выполняет УС. Также стоит отметить, что в процессе управления и в том, и в другом случае присутствуют несколько условий:

- существование возможной реакции интегрируемой ИС на входные данные;
- влияние полученной реакции на состояние PLM системы (возможность добавления данных в комплекс без ошибок и рисков для всего производственного комплекса).

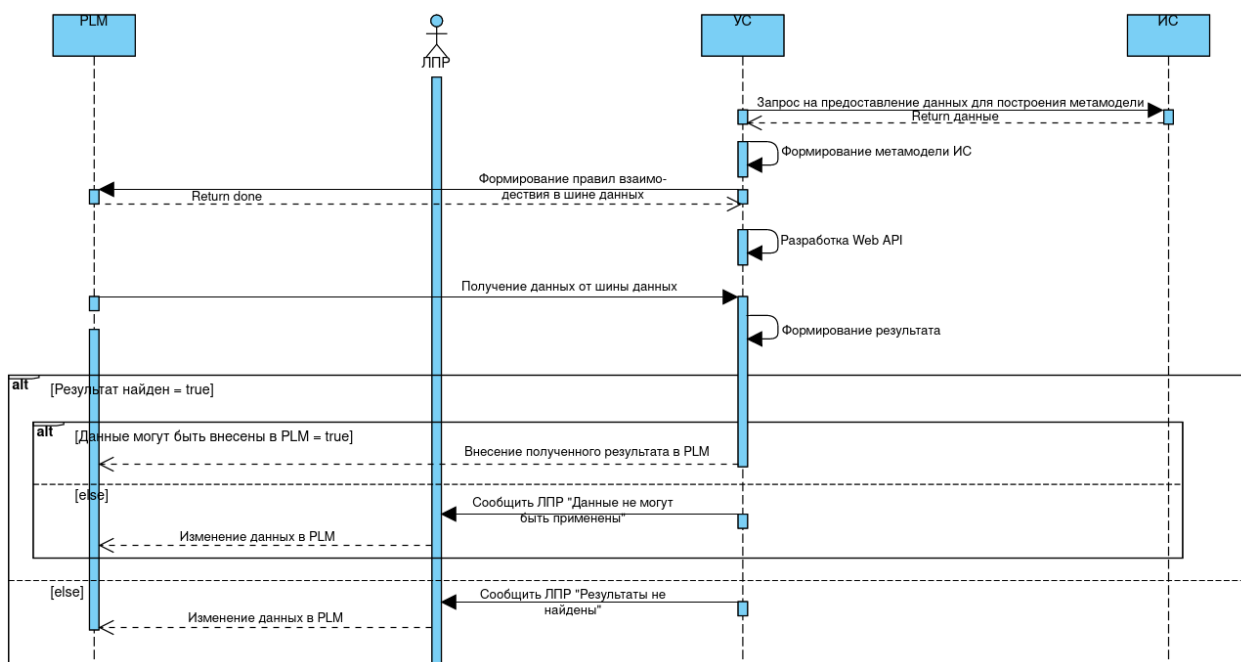


Рис. 9. Процесс организации взаимодействия и поддержания системы в работоспособном состоянии с участием УС и ЛПП

Таким образом, можно увидеть, что основная задача ЛПП состоит в корректировке данных общего комплекса PLM, не взаимодействуя при этом с интегрируемой информационной системой. Общий процент сценариев, где должен участвовать человек при работе с системой без УС, составляет 85%, а общее количество сценариев при работе с системой с УС составляет 18%. Основной задачей является, в случае использования УС, реакция на ошибки при формировании результатов работы интегрируемой ИС. Получаем снижение роли лица, принимающего решения, на 67%.

Заключение. В статье приведен подход к управлению на основе данных применительно к производственным процессам. Подход подразумевает использование метамодели данных информационной системы и иерархическую базу правил поведения информационной системы с нечетким логическим выводом. Приведены основные выводы, подтверждающие целесообразность использования данного подхода в задачах отображения данных информационных систем и управления информационной системой с целью снижения человеческого фактора при организации управления сторонними системами на производстве. Задачи являются актуальными на многих производствах Российской Федерации, в том числе на производственных предприятиях Ульяновской области.

Благодарности. Исследование выполнено в рамках государственного задания № 075-03-2023-143 по проекту «Исследование интеллектуальной предиктивной аналитики на базе интеграции методов конструирования признаков гетерогенных динамических данных для машинного обучения и методов предиктивного мультимодального анализа данных».

Список источников

1. Terzi S., Bouras A., Dutta D., Garetti M., Kiritsis D. Product lifecycle management – from its history to its new role. International journal product lifecycle management, 2010, vol. 4, no. 4, pp. 360–389.
2. Menge F. Enterprise service bus. Free and open source software conference. 2007, available at: https://programm.froscon.org/2007/attachments/15-falko_menge_-_enterprise_service_bus.pdf (accessed: 05/30/2024)
3. Шведенко В. Н. Архитектура интегрированной информационной системы, обеспечивающая свойство поведения / В.Н. Шведенко, О.В. Щекочихин // Научно-технический вестник информационных технологий, механики и оптики, 2016. – Т. 16. – №. 6. – С. 1078-1083.
4. Чекал Е.Г. Надежность информационных систем: учебное пособие : в 2 ч. Ч. 1 / Е.Г. Чекал, А.А. Чичев. – Ульяновск: УлГУ, 2012. – 118 с.

5. Feng G. A survey on analysis and design of model-based fuzzy control systems. *IEEE Transactions on Fuzzy systems*, 2006, vol. 14, no. 5, pp. 676-697.
6. Sudarsan R., Fenves S.J., Sriram R.D., Wang F. A product information modeling framework for product lifecycle management. *Computer-Aided design* 37, 2005, pp. 1399–1411
7. Alahi MEE, Sukkuea A., Tina F.W. et al. Integration of IoT-enabled technologies and Artificial Intelligence (AI) for Smart City scenario: recent advancements and future trends. *Sensors*. 2023, vol. 23, no. 11, DOI:10.3390/s23115206.
8. Sacks R., Brilakis I., Pikas E., et al. Construction with digital twin information systems. *Data-Centric engineering*, 2020, vol. 1:e14, DOI:10.1017/dce.2020.16.
9. Cimini Ch., Pirola F, Pinto R., Cavalieri S., A human-in-the-loop manufacturing control architecture for the next generation of production systems. *Journal of manufacturing systems*, 2020, vol. 54, pp. 258-271.
10. Hou Z.S., Xu J.X. On Data-driven control theory: the state of the art and perspective. *Acta automatica sinica*, 2009, vol. 35, no. 6, pp. 650–667.
11. Mamdani E.H. Application of fuzzy algorithms for control of simple dynamic plant. In: *Proceedings of the institution of electrical engineers*, 1974, vol. 121, pp. 1585–1588.
12. Gonzalez A., Perez R., Verdegay J.L. Learning the structure of a fuzzy rule: a genetic approach. *Fuzzy syst. Artif. Intell.*, 1994, vol. 3, no. 1, pp. 57–70.
13. Duckstein, L., et al. *Fuzzy rule-based modeling with applications to geophysical, biological, and engineering systems*. CRC Press, Boca Raton, 1995, vol. 8, DOI:10.1201/9780138755133.
14. Chi Z., Yan H., Pham T. *Fuzzy algorithms: with applications to image processing and pattern recognition*. World scientific, Singapore, 1996, vol. 10.
15. Razak T.R., Fauzi S.S.M., Gining R.A.J. et al. Hierarchical fuzzy systems: interpretability and complexity. *Indones. J. Electr. Eng. Inform.*, 2021, vol. 9, no. 2, pp. 478–489.
16. Zhang X., Onieva E., Perallos A., et al. Hierarchical fuzzy rule-based system optimized with genetic algorithms for short term traffic congestion prediction. *Transport. Res. C: Emerg. Technol*, 2014, vol. 43, pp. 127–142.
17. Duan J.-C., Chung F.-L. Cascaded fuzzy neural network model based on syllogistic fuzzy reasoning. *IEEE trans. fuzzy syst.*, 2001, vol. 9, no. 2, pp. 293–306.
18. Jang J.-S.R. Fuzzy modeling using generalized neural networks and Kalman filter algorithm. *AAAI* 91, 1991, pp. 762–767.
19. Комарцова Л.Г. Эволюционные методы формирования нечетких баз правил / Л. Г. Комарцова // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2011): материалы международной научно-технической конференции, Минск, 10-12 февраля 2011 г. / Белорусский государственный университет информатики и радиоэлектроники; редкол.: В. В. Голенков (отв. ред.) [и др.]. – Минск, 2011. – С. 181–184. – URL: <https://libeldoc.bsuir.by/handle/123456789/4495> (дата обращения: 30.05.2024).
20. Камалетдинова Л.Р. Моделирование управляющей системы / Л.Р. Камалетдинова А.А. Романов. // *System analysis & mathematical modeling*, 2024. – Т. 6. – № 1. – С. 60–77. – DOI:10.17150/2713-1734.2024.6(1).60-77. – EDN UYCUWC. – URL: <http://samm-bgu.ru/reader/article.aspx?id=26482> (дата обращения: 30.05.2024).
21. Продукт службы поддержки Faveo Service Desk – URL: <https://www.faveohelpdesk.com/servicedesk/?lang=ru> (дата обращения: 30.05.2024).
22. Kamaletdinova L., Romanov A., Yarushkina N. Industrial data-driven control, 2024 International Russian smart industry conference (SmartIndustryCon), Sochi, Russian Federation, 2024, pp. 62-67.

Камалетдинова Лилия Рашидовна. Аспирант направления Компьютерное моделирование и автоматизация проектирования, ассистент кафедры Информационные системы УлГТУ. Область интересов: моделирование информационных систем и предиктивная аналитика. ORCID 0009-0006-2929-9757, SPIN: 8878-6567, AuthorID: 1222414, lribragimova1995@gmail.com, 432027, г. Ульяновск, ул. Северный Венец, д. 32.

Романов Антон Алексеевич. Кандидат технических наук, доцент кафедры Информационные системы УлГТУ, заведующий кафедрой Информационные системы УлГТУ. Область интересов: интеллектуальная система хранения и обработки информации. ORCID 0000-0001-5275-7628, SPIN: 4990-5739, AuthorID: 684949, romanov73@gmail.com, 432027, г. Ульяновск, ул. Северный Венец, д. 32.

Филиппов Алексей Александрович. Кандидат технических наук, доцент кафедры Информационные системы УлГТУ. Область интересов: онтологическое моделирование и интеллектуальный анализ данных. ORCID 0000-0003-0008-5035, SPIN: 2106-1545, AuthorID: 708454, al.filippov@ulstu.ru, 432027, г. Ульяновск, ул. Северный Венец, д. 32.

Ярушкина Надежда Глебовна. Доктор технических наук, профессор кафедры Информационные системы УлГТУ, ректор УлГТУ. Область интересов: мягкие вычисления, нечеткая логика, гибридные системы. ORCID 0000-0002-5718-8732, SPIN: 1658-2640, AuthorID: 10358, jng@ulstu.ru, 432027, г. Ульяновск, ул. Северный Венец, д. 32.

UDC 004.896

DOI:10.25729/ESI.2024.35.3.011

Rule base forming for PLM systems**Liliya R. Kamaletdinova, Anton A. Romanov, Aleksey A. Filippov, Nadezhda G. Yarushkina**

Ulyanovsk state technical university,

Russia, Ulyanovsk, *lribragimova1995@gmail.com*

Abstract. There is a need to integrate a PLM system and a third-party information system. The information system is not part of the PLM complex, but is involved in solving problems of information support for managing production processes. At the moment, the task of such integration is performed by the analyst. It will form a structural and process model of an integrated information system. Based on the model, rules for interaction with the system are formed. The operator and the decision maker (DM) are involved in the operation process.

An approach is proposed to reduce the load on both the analyst and the operator and decision maker.

Keywords: Rule-based systems, rule base, PLM, data-driven control

Acknowledgements: This study was supported the Ministry of Science and Higher Education of Russia in framework of project No. 075-03-2023-143 "The study of intelligent predictive analytics based on the integration of methods for constructing features of heterogeneous dynamic data for machine learning and methods of predictive multimodal data analysis".

References

1. Terzi S., Bouras A., Dutta D., Garetti M., Kiritsis D. Product lifecycle management – from its history to its new role. *International journal product lifecycle management*, 2010, vol. 4, no. 4, pp. 360–389.
2. Menge F. Enterprise service bus. Free and open source software conference. 2007, available at: https://programm.froscon.org/2007/attachments/15-falko_menge_-_enterprise_service_bus.pdf (accessed: 05/30/2024)
3. Shvedenko V.N., Shchekochikhin O.V. Arkhitektura integrirovannoy informatsionnoy sistemy, obespechivayushchaya svoystvo povedeniya [Architecture of an integrated information system that ensures behavioral properties]. *Nauchno-tekhnicheskiiy vestnik informatsionnykh tekhnologiy, mekhaniki i optiki* [Scientific and technical bulletin of information technologies, mechanics and optics], 2016, vol. 16, no. 6, pp. 1078–1083.
4. Chekal E.G., Chichev A.A. Nadezhnost' informatsionnykh sistem: uchebnoye posobiye: v 2 ch. CH. 1 [Reliability of information systems: textbook: in 2 parts. Part 1]. Ulyanovsk, UIGU, 2012, 118 p.
5. Feng G. A survey on analysis and design of model-based fuzzy control systems. *IEEE Transactions on Fuzzy systems*, 2006, vol. 14, no. 5, pp. 676–697.
6. Sudarsan R., Fenves S.J., Sriram R.D., Wang F. A product information modeling framework for product lifecycle management. *Computer-Aided design* 37, 2005, pp. 1399–1411
7. Alahi MEE, Sukkuea A., Tina F.W. et al. Integration of IoT-enabled technologies and Artificial Intelligence (AI) for Smart City scenario: recent advancements and future trends. *Sensors*. 2023, vol. 23, no. 11, DOI:10.3390/s23115206.
8. Sacks R., Brilakis I., Pikas E., et al. Construction with digital twin information systems. *Data-Centric engineering*, 2020, vol. 1:e14, DOI:10.1017/dce.2020.16.
9. Cimini Ch., Pirola F, Pinto R., Cavalieri S., A human-in-the-loop manufacturing control architecture for the next generation of production systems. *Journal of manufacturing systems*, 2020, vol. 54, pp. 258–271.
10. Hou Z.S., Xu J.X. On Data-driven control theory: the state of the art and perspective. *Acta automatica sinica*, 2009, vol. 35, no. 6, pp. 650–667.
11. Mamdani E.H. Application of fuzzy algorithms for control of simple dynamic plant. In: *Proceedings of the institution of electrical engineers*, 1974, vol. 121, pp. 1585–1588.
12. Gonzalez A., Perez R., Verdegay J.L. Learning the structure of a fuzzy rule: a genetic approach. *Fuzzy syst. Artif. Intell.*, 1994, vol. 3, no. 1, pp. 57–70.
13. Duckstein, L., et al. Fuzzy rule-based modeling with applications to geophysical, biological, and engineering systems. CRC Press, Boca Raton, 1995, vol. 8, DOI:10.1201/9780138755133.
14. Chi Z., Yan H., Pham T. Fuzzy algorithms: with applications to image processing and pattern recognition. World scientific, Singapore, 1996, vol. 10.
15. Razak T.R., Fauzi S.S.M., Gining R.A.J. et al. Hierarchical fuzzy systems: interpretability and complexity. *Indones. J. Electr. Eng. Inform.*, 2021, vol. 9, no. 2, pp. 478–489.
16. Zhang X., Onieva E., Perallos A., et al. Hierarchical fuzzy rule-based system optimized with genetic algorithms for short term traffic congestion prediction. *Transport. Res. C: Emerg. Technol*, 2014, vol. 43, pp. 127–142.
17. Duan J.-C., Chung F.-L. Cascaded fuzzy neural network model based on syllogistic fuzzy reasoning. *IEEE trans. fuzzy syst.*, 2001, vol. 9, no. 2, pp. 293–306.

18. Jang J.-S.R. Fuzzy modeling using generalized neural networks and Kalman filter algorithm. AAAI 91, 1991, pp. 762–767.
19. Komartsova L.G. Evolyutsionnyye metody formirovaniya nechetkikh baz pravil [Evolutionary methods for forming fuzzy rule bases]. Otkrytyye semanticheskiye tekhnologii proyektirovaniya intellektual'nykh sistem (OSTIS-2011): materialy mezhdunarodnoy nauchno-tekhnicheskoy konferentsii, Minsk, 10-12 fevralya 2011 g. Belorusskiy gosudarstvennyy universitet informatiki i radioelektroniki; redkol.: V. V. Golenkov. [Open semantic technologies for intelligent systems (OSTIS-2011): Materials of the international scientific and technical conference, Minsk, 10-12 february 2011. Belarusian state university of informatics and radioelectronics; editor: V.V. Golenkov], Minsk, 2011, pp. 181–184, available at: <https://libeloc.bsuir.by/handle/123456789/4495> (accessed: 05/30/2024).
20. Kamaletdinova, L., Romanov A. Modelirovaniye upravlyayushchey sistemy [Control system design]. System analysis & mathematical modeling, 2024, vol. 6., pp. 60-77, DOI:10.17150/2713-1734.2024.6(1).60-77, EDN UYCUWC, available at: <http://samm-bgu.ru/reader/article.aspx?id=26482> (дата обращения: 30.05.2024).
21. Faveo Service Desk. Available at: <https://www.faveohelpdesk.com/servicedesk> (accessed: 05/30/2024).
22. Kamaletdinova L., Romanov A., Yarushkina N. Industrial data-driven control, 2024 International Russian smart industry conference (SmartIndustryCon), Sochi, Russian Federation, 2024, pp. 62-67.

Kamaletdinova Lilia Rashidovna. Postgraduate student majoring in Computer modeling and design automation, assistant at the Information systems department. She has articles in the field of information systems modeling and predictive analytics. ORCID 0009-0006-2929-9757, SPIN: 8878-6567, AuthorID: 1222414, lribragimova1995@gmail.com, 432027, Russia, Ulyanovsk, Severniy Venec, 32.

Romanov Anton Alekseevich. Candidate of technical sciences, associate professor of the Information systems department at Ulyanovsk state technical university, Head of the information systems department at Ulyanovsk state technical university. Area of interest: intelligent systems for storing and processing information. ORCID 0000-0001-5275-7628, SPIN: 4990-5739, AuthorID: 684949, romanov73@gmail.com, 432027, Russia, Ulyanovsk, Severniy Venec, 32.

Filippov Aleksey Aleksandrovich. Candidate of technical sciences, associate professor of the Information systems department at Ulyanovsk state technical university. Area of interest: ontological modeling and data mining. ORCID 0000-0003-0008-5035, SPIN: 2106-1545, AuthorID: 708454, al.filippov@ulstu.ru, 432027, Russia, Ulyanovsk, Severniy Venec, 32.

Yarushkina Nadezhda Glebovna. Doctor of technical sciences, professor of the Information systems department at Ulyanovsk state technical university, rector of Ulyanovsk state technical university. Area of interest: soft computing, fuzzy logic, hybrid systems. ORCID 0000-0002-5718-8732, SPIN: 1658-2640, AuthorID: 10358, jng@ulstu.ru, 432027, Russia, Ulyanovsk, Severniy Venec, 32.

Статья поступила в редакцию 31.05.2024; одобрена после рецензирования 01.10.2024; принята к публикации 10.10.2024.

The article was submitted 05/31/2024; approved after reviewing 10/01/2024; accepted for publication 10/10/2024.

Математическое моделирование и его применение в научных исследованиях

УДК 519.673

DOI:10.25729/ESI.2024.35.3.012

Модификация алгоритма минимальной степени для повышения эффективности вычислений при итерационном решении контактной задачи теории упругости методом конечных элементов

Дудаев Михаил Алексеевич

Иркутский государственный университет путей сообщения,
Россия, Иркутск, dudaev_ma@mail.ru

Аннотация. В статье предложена модификация классического алгоритма минимальной степени, предназначенного для упорядочения структуры глобальной матрицы коэффициентов системы линейных алгебраических уравнений метода конечных элементов, с целью повышения эффективности вычислительного алгоритма решения контактной задачи теории упругости при моделировании сопряжений деталей специальными контактными элементами сопряжения конструкций, применение которых обуславливает итерационное решение задачи с постепенным уточнением их состояния. Предложенная модификация реализована совместно с прямым методом Холецкого решения систем линейных алгебраических уравнений, вычислительная эффективность которого проявляется в наибольшей степени при решении систем с симметричной, положительно определенной матрицей коэффициентов, преимущественно имеющей в методе конечных элементов, разреженную структуру. Важной проблемой в методе Холецкого является заполнение «треугольного множителя» по сравнению с исходной матрицей в процессе разложения; уменьшение заполнения реализуется обычно с использованием известных эвристических алгоритмов, среди которых существенное распространение получил алгоритм минимальной степени. Модификация этого алгоритма, предложенная в работе, симметричной перестановкой строк и столбцов упорядочивает структуру исходной глобальной матрицы таким образом, что позволяет в процессе итерационного решения контактных задач теории упругости, проводить частичное разложение Холецкого на повторных итерациях с экономией времени расчета по сравнению с полным разложением. В работе показано, что структура матрицы, упорядоченная модифицированным алгоритмом, в целом, является менее предпочтительной, нежели при использовании классического алгоритма минимальной степени, так как претерпевает большее заполнение, однако, эффективность применения вычислительного алгоритма достигается с проведением повторных итераций при постепенном уточнении состояния контактных конечных элементов сопряжения конструкций. С использованием тестовой конечноэлементной модели сборного ротора, а также модели роторной системы реального авиационного газотурбинного двигателя, определены и показаны портреты исходных матриц и их треугольных множителей, проведены оценки количества арифметических операций и условного времени их выполнения при использовании классического и модифицированного алгоритма минимальной степени.

Ключевые слова: метод конечных элементов, СЛАУ, контактная задача, метод Холецкого, разложение, факторизация, алгоритм минимальной степени, скалярные произведения, заполнение

Цитирование: Дудаев М.А. Модификация алгоритма минимальной степени для повышения эффективности вычислений при итерационном решении контактной задачи теории упругости методом конечных элементов / М.А. Дудаев // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С.133-147. – DOI: 10.25729/ESI.2024.35.3.012.

Введение. При решении задач механики твердого деформируемого тела широко используется метод конечных элементов (МКЭ) [1]. Его применение в задачах статики и динамики во многом связано с построением симметричных, положительно определенных и разреженных матриц [2] (жесткости, масс и демпфирования), являющихся матрицами коэффициентов в разрешающих системах линейных алгебраических уравнений (СЛАУ). Хранение таких матриц в памяти ЭВМ обычно реализуется с использованием компактных схем хранения [3].

Численное решение СЛАУ в МКЭ в задачах статики и динамики обычно выполняется с использованием прямых методов [4], выполняющих его с доступной машинной точностью за конечное и предсказуемое количество арифметических операций. Наибольшее распространение при решении СЛАУ с положительно определенными симметричными разреженными матрицами получил метод Холецкого [4-6], обеспечивающий решение за минимальное количество арифметических операций. В этом случае матрица коэффициентов $[A]$ представляется в виде:

$$[A] = [L] \cdot [L]^T, \quad (1)$$

где $[L]$ – нижнетреугольная матрица, называемая треугольным «множителем» Холецкого, а процедура ее вычисления – разложением (факторизацией).

В работе [5] исследованы несколько алгоритмов вычисления треугольного множителя, однако, при использовании компактных схем хранения матриц $[A]$ и $[L]$ преимущество имеет алгоритм в форме скалярных произведений. В таком случае элементы l треугольного множителя Холецкого $[L]$ определяются выражениями:

$$l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}, i = 1 \dots n; l_{ji} = \frac{1}{l_{ii}} \left(a_{ji} - \sum_{k=1}^{i-1} l_{ik} \cdot l_{jk} \right), j = i + 1 \dots n, \quad (2)$$

где i и j – номер строки и столбца элемента соответственно; n – количество строк матриц $[A]$ и $[L]$; a – элемент матрицы $[A]$.

При выполнении первичного присваивания $[L] = [A]$ порядок обращения к элементам матрицы $[L]$ в процессе ее разложения в форме скалярных произведений показан на рис. 1.

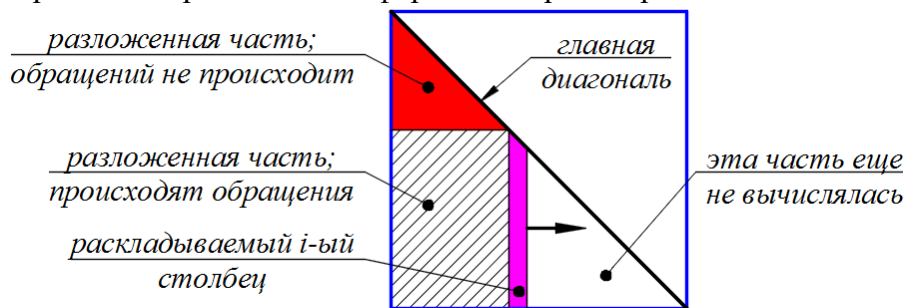


Рис. 1. Порядок обращения к элементам матрицы $[L]$ в процессе ее разложения

Недостатком прямых методов решения СЛАУ является «заполнение» матрицы [7], при котором в процессе разложения в матрице $[L]$ появляются дополнительные ненулевые значения. Эта особенность при использовании компактных схем хранения в памяти ЭВМ не позволяет хранить треугольный множитель $[L]$ в массивах, выделенных под исходную матрицу $[A]$, а также увеличивает время решения СЛАУ из-за дополнительных арифметических операций.

Заполнение треугольного множителя зависит от структуры исходной матрицы $[A]$, которую можно изменить симметричной перестановкой строк и столбцов [4, 5], что в математической интерпретации означает умножение исходной матрицы $[A]$ слева и справа на матрицу перестановок $[P]$:

$$[B] = [P]^T \cdot [A] \cdot [P] \quad (3)$$

В матрице перестановок $[P]$ в каждой строке присутствует всего один ненулевой элемент, равный единице; умножение слева на матрицу $[P]^T$ обеспечивает перестановку строк исходной матрицы $[A]$, а умножение справа на матрицу $[P]$ – перестановку столбцов.

Для уменьшения заполнения в треугольном множителе обычно применяются эвристические алгоритмы «минимальной степени» (АМС), «вложенных сечений», обратный алгоритм «Катхилла-Макки» и др., а также их комбинации [4, 5].

Классический алгоритм минимальной степени. При анализе динамического поведения многокаскадных сборных роторных систем турбомашин [8] с применением специально разработанного комплекса программ [9], реализованного на языке программирования FORTRAN 90 [10], используется АМС, показавший в этой задаче более эффективный результат уменьшения заполнения треугольного множителя $[L]$, и, как следствие, позволяющий существенно экономить память ЭВМ [4], а также получивший наибольшее распространение [5].

Принцип работы алгоритма поясняется на примере конечноэлементной конструкции, показанной на рис. 2а, с использованием элементов теории графов [11, 12]. Соответствующий исходной структуре матрицы граф смежности показан на рис. 2б, где черным цветом указаны номера вершин графа (узлов КЭ), а красным – степени вершин.

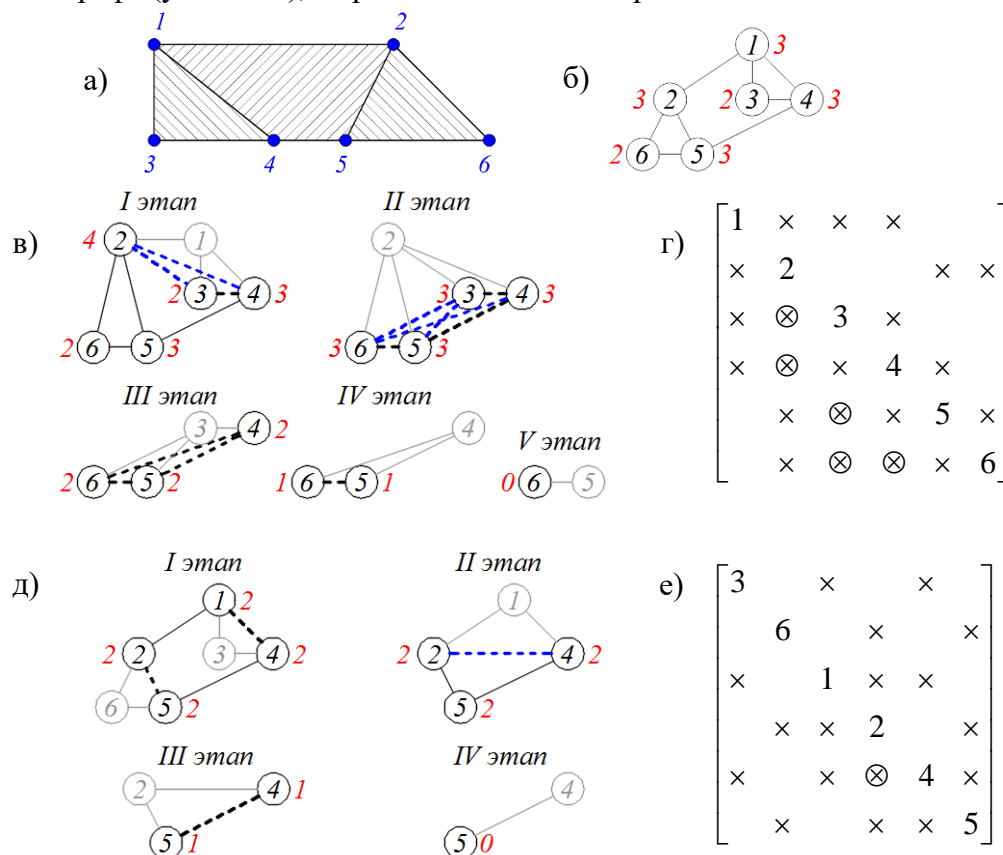


Рис. 2. Исключение вершин графа смежности в разной очередности при его ассоциации с узлами конечноэлементной модели

Порядок исключения вершин графа **без перестановки строк и столбцов** показан на рис. 2в. После исключения вершины X_i добавляются ребра («клики») так, чтобы все вершины из множества $Adj(X_i)$ оказались попарно смежными. Случай совпадения добавленного ребра с уже существующим свидетельствует о проведении математических операций с числом в процессе разложения, а случай появления новых ребер – о заполнении. Структура (портрет) матрицы жесткости при использовании начальных номеров узлов в качестве номеров строк и столбцов показана на рис. 2г, где ненулевые внедиагональные элементы исходной матрицы обозначены символом « \times », а поддиагональные заполнения, получаемые в ходе разложения – « \otimes ». В соответствии с рис. 2г, без перестановки строк после разложения в матрице появляется 5 новых элементов под главной диагональю.

На рис. 2д показано исключение вершин **в соответствии с АМС**. После предварительного расчета степеней всех вершин производится исключение вершин 3 и 6, имеющих минимальную степень, равную двум. После их исключения добавляются ребра 2-5 и 1-4, которые совпадают с уже существующими ребрами. Степени оставшихся вершин пересчитываются и

в приведенном примере оказываются равными. В следующую очередь происходит исключение вершины 1 и добавляется ребро 2-4 (клика), поскольку каждая из этих вершин являлась смежной с исключенной, что свидетельствует о заполнении элемента l_{42} . Процесс повторяется до исключения всех вершин. Структура матрицы, совмещенная с множителем Холецкого после упорядочения АМС показана на рис. 2, е.

Особенности решения контактной задачи теории упругости прямыми методами в МКЭ. При решении контактной задачи анализа динамического поведения многокаскадных сборных роторных систем турбомашин, математическая модель которой представлена в работах [13-15], используется специальный контактный конечный элемент сопряжения конструкций – КЭСК (рис. 3).

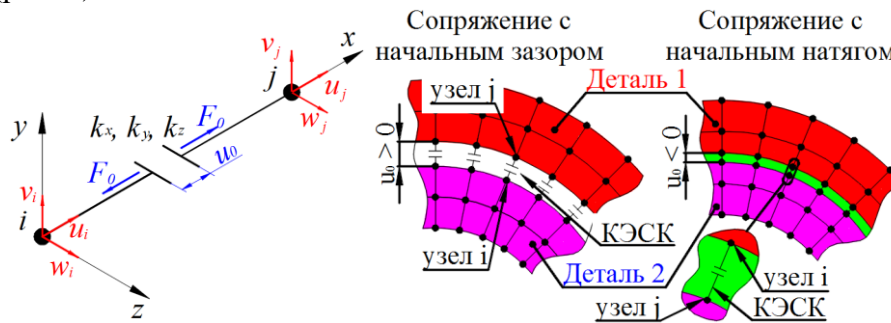


Рис. 3. Контактный конечный элемент сопряжения конструкций

Данный КЭ может находиться в одном из двух состояний: «открытом» и «закрытом». Оценка состояний КЭСК проводится на основании невязки перемещений Δu его узлов i и j :

$$\Delta u = u_i - u_j + u_0, \quad (4)$$

где u_i и u_j – перемещения узлов i и j КЭСК соответственно в направлении оси x локальной системы координат; u_0 – начальный зазор/натяг КЭСК.

При условии $\Delta u > 0$ КЭСК находится в открытом состоянии, а при $\Delta u \leq 0$ – в закрытом.

В зависимости от состояния КЭСК изменяется его матрица жесткости и вектор контактных сил:

$$[K]_c^+ = \begin{bmatrix} [K]^+ & -[K]^+ \\ -[K]^+ & [K]^+ \end{bmatrix}; [K]^+ = \begin{bmatrix} k_x^+ & 0 & 0 \\ 0 & k_y^+ & 0 \\ 0 & 0 & k_z^+ \end{bmatrix}; \{F\}_c^+ = 0; \quad (5)$$

$$[K]_c^- = \begin{bmatrix} [K]^- & -[K]^- \\ -[K]^- & [K]^- \end{bmatrix}; [K]^- = \begin{bmatrix} k_x^- & 0 & 0 \\ 0 & k_y^- & 0 \\ 0 & 0 & k_z^- \end{bmatrix}; \{F\}_c^- = k_x^- \cdot \Delta u \begin{Bmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{Bmatrix}, \quad (6)$$

где $[K]_c^+$ и $[K]_c^-$ – матрицы жесткости КЭСК в открытом и закрытом состоянии соответственно; k_x^+ и k_x^- – коэффициенты продольной жесткости КЭСК в открытом и закрытом состоянии соответственно [15]; k_y^+ , k_z^+ и k_y^- , k_z^- – коэффициенты поперечной жесткости КЭСК в открытом и закрытом состоянии соответственно; $\{F\}_c^+$ и $\{F\}_c^-$ – вектор сил КЭСК в открытом и закрытом состоянии соответственно.

Смена состояний КЭСК в процессе статического или динамического анализа обуславливает итерационное решение контактной задачи с постепенным уточнением на каждой итерации состояния всех КЭСК. Блок-схема работы алгоритма решения контактной задачи МКЭ в статической постановке представлена на рис. 4.

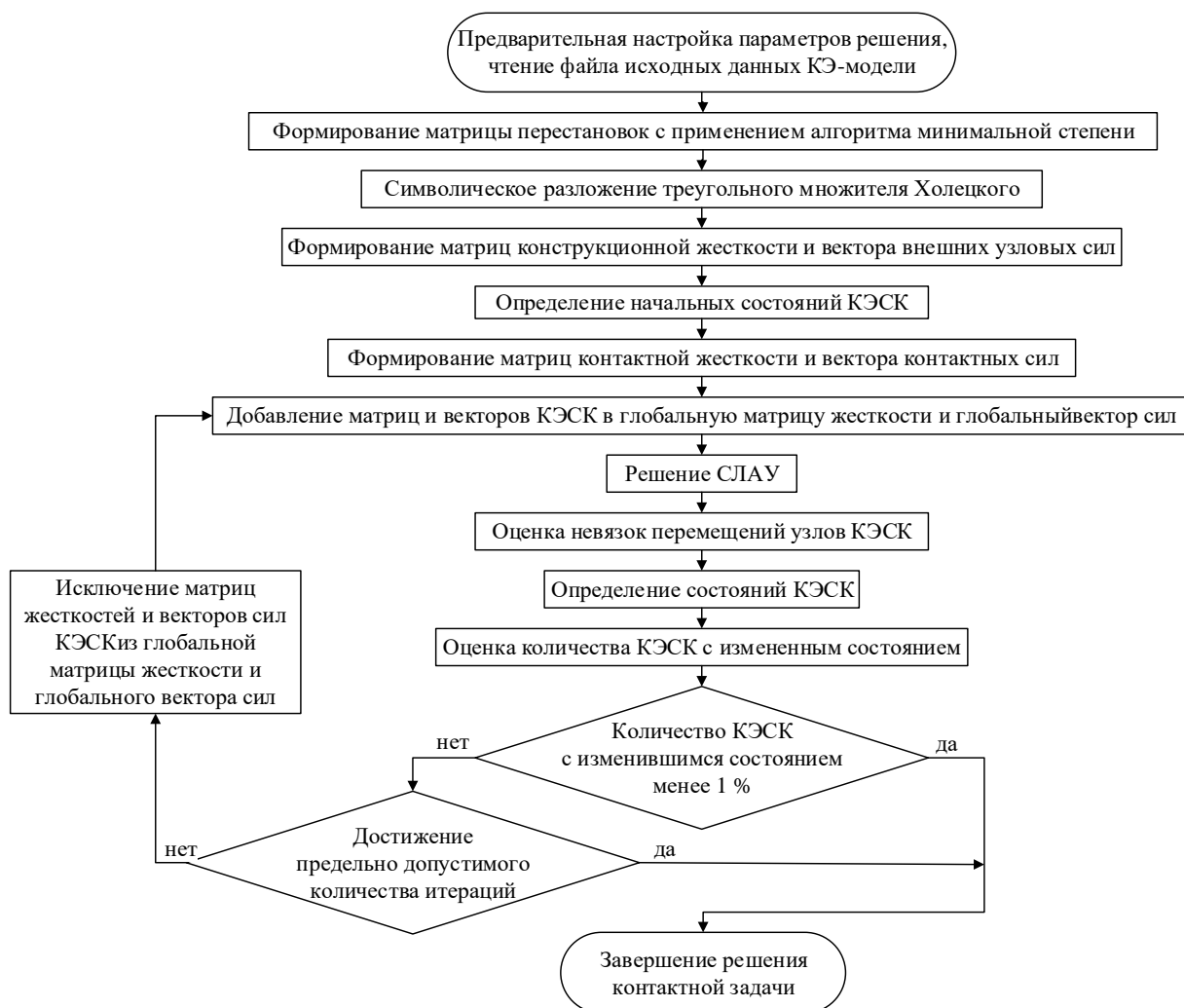


Рис. 4. Блок-схема алгоритма решения контактной задачи теории упругости методом конечных элементов в статической постановке

Основную трудоемкость вычислительного алгоритма при итерационном решении контактной задачи составляет решение СЛАУ, а именно вещественное разложение Холецкого, проводимое на каждой итерации ввиду изменений в глобальной матрице жесткости из-за смены состояний КЭСК.

Для демонстрации указанной особенности рассмотрен пример, показанный на рис. 5 в виде системы из пяти стержневых конечных элементов (КЭ) с номерами 1, 2, 4-6 (номера КЭ обведены окружностями) с одинаковой продольной жесткостью k и одним КЭСК (№ 3) с осевой жесткостью k_k , величина которой изменяется в зависимости от его состояния.

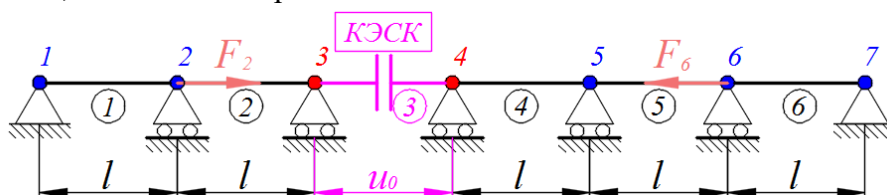


Рис. 5. Конечнoэлементная модель стержневой системы с одним КЭСК

Матрица жесткости представленной системы с учетом закрепления узлов 1 и 7 в горизонтальном направлении (номера узлов показаны на рис. 5) имеет вид [1] (строки № 1 и 7 матрицы вычеркнуты):

$$[K] = \begin{bmatrix} 2k & -k & 0 & 0 & 0 \\ -k & k + k_K & -k_K & 0 & 0 \\ 0 & -k_K & k + k_K & -k & 0 \\ 0 & 0 & -k & 2k & -k \\ 0 & 0 & 0 & -k & 2k \end{bmatrix}. \quad (7)$$

Треугольный множитель Холецкого [L] для матрицы [K] имеет вид

$$[L] = \begin{bmatrix} \sqrt{2k} & 0 & 0 & 0 & 0 \\ \frac{-k}{\sqrt{2k}} & \sqrt{\frac{k}{2} + k_K} & 0 & 0 & 0 \\ 0 & -\sqrt{\frac{2k_K^2}{k + 2k_K}} & \sqrt{\frac{k(k + 3k_K)}{k + 2k_K}} & 0 & 0 \\ 0 & 0 & -\sqrt{\frac{k(k + 2k_K)}{k + 3k_K}} & \sqrt{\frac{k(k + 4k_K)}{k + 3k_K}} & 0 \\ 0 & 0 & 0 & -\sqrt{\frac{k(k + 3k_K)}{k + 4k_K}} & \sqrt{\frac{k(k + 5k_K)}{k + 4k_K}} \end{bmatrix}. \quad (8)$$

В представленном примере видно, что изменение величины контактной жесткости k_K повлечет за собой изменение величин элементов треугольного множителя, начиная с l_{22} и расположенных ниже и правее него под главной диагональю, поскольку во все эти элементы входит величина k_K . Таким образом, даже в примере с одним КЭСК изменение его состояния приведет к необходимости практически полного повторного разложения Холецкого. При наличии множества КЭСК их состояния уточняются в процессе итерационного решения, поэтому разложения Холецкого будут вычисляться на каждой итерации, что существенно сказывается на времени расчета даже в статической постановке задачи. В задачах динамики [1] решение контактной задачи проводится на каждом шаге интегрирования по времени [13-15], что может неприемлемо увеличивать время расчета КЭ модели, особенно, на этапах ее отладки, когда решение задачи требуется проводить многократно.

Модификация алгоритма минимальной степени. В процессе отладки итерационного решения контактной задачи теории упругости МКЭ при анализе динамического поведения многокаскадных сборных роторных систем турбомашин был принят во внимание тот факт, что изменения коэффициентов жесткостей на каждой итерации происходят не во всех строках и столбцах глобальной матрицы жесткости, а только тех, в которые входят коэффициенты жесткости КЭСК. В связи с тем, что для упорядочения структуры глобальной матрицы жесткости уже был программно реализован АМС, возникла идея его модификации таким образом, чтобы при проведении упорядочения строки и столбцы, связанные с узлами КЭСК, симметрично переставлялись в конец матрицы (вниз и вправо). Для этого к реальным степеням узлов (соответствующим вершинам графа смежности), связанных с КЭСК, достаточно прибавить большое число, например, соответствующее размерности n матрицы. Это приводит к тому, что степени таких вершин графа смежности (узлов КЭСК) становятся больше, чем у остальных вершин. В таком случае продолжение работы АМС приводит к «принудительной» симметричной перестановке строк и столбцов матрицы, связанных с узлами КЭСК в ее конец.

Работа модифицированного таким образом АМС рассматривается на уже представленном примере (рис. 5). Соответствующий КЭ модели граф смежности показан на рис. 6а.

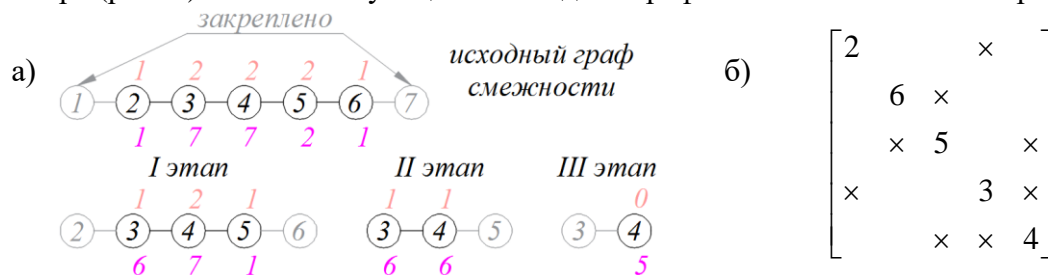


Рис. 6. Исключение вершин графа смежности с использованием модифицированного алгоритма минимальной степени

Узлы 1 и 7 в КЭ модели закреплены, поэтому соответствующие им строки вычеркнуты из матрицы, а соответствующие узлам вершины – исключены из графа смежности до начала работы АМС. На первой итерации АМС определяются естественные степени вершин графа, показанные на рисунке над вершинами красным цветом. После этого модифицированным АМС определяются вершины, ассоциированные с узлами КЭСК, и к степеням этих вершин прибавляется размерность матрицы (равная пяти) так, чтобы полученные величины степеней были самыми большими в графе. Модифицированные степени показаны под графом смежности малиновым цветом. Затем, в соответствии с работой АМС, из графа исключаются вершины 2 и 6, степень которых минимальна и равна единице. На второй итерации модифицированного АМС пересчитываются реальные степени оставшихся вершин; к степеням вершин, ассоциированных с узлами КЭСК, прибавляется размерность матрицы; после этого исключается вершина 5, степень которой минимальна и равна единице. Аналогично проводится третья итерация, в ходе которой исключаются последние вершины 3 и 4. Портрет матрицы после упорядочения модифицированным АМС показан на рис. 6б.

При разложении Холецкого представленная в примере матрица не претерпевает изменений как в исходном, так и в упорядоченном варианте, что обусловлено линейной структурой графа смежности. Однако в преобразованном виде матрица имеет вид:

$$[K] = \begin{bmatrix} 2k & 0 & 0 & -k & 0 \\ 0 & 2k & -k & 0 & 0 \\ 0 & -k & 2k & 0 & -k \\ -k & 0 & 0 & k + k_K & -k_K \\ 0 & 0 & -k & -k_K & k + k_K \end{bmatrix}, \quad (9)$$

а ее треугольный множитель Холецкого $[L]$:

$$[L] = \begin{bmatrix} \sqrt{2k} & 0 & 0 & 0 & 0 \\ 0 & \sqrt{2k} & 0 & 0 & 0 \\ 0 & -\sqrt{\frac{k}{2}} & \sqrt{\frac{3}{2}}k & 0 & 0 \\ \frac{-k}{\sqrt{2k}} & 0 & 0 & \sqrt{\frac{1}{2}}k + k_K & 0 \\ 0 & 0 & -\sqrt{\frac{2}{3}}k & -\sqrt{\frac{2k_K^2}{k + 2k_K}} & \sqrt{\frac{k(k + 5k_K)}{3(k + 2k_K)}} \end{bmatrix}. \quad (10)$$

В таком представлении видно, что жесткости контактных элементов встречаются в матрице и треугольном множителе Холецкого, начиная только с элемента 44 (в соответствии с номерами строк и столбцов матрицы) и далее (снизу и справа от этой позиции).

Это обстоятельство, а также порядок обращения к элементам матрицы при ее разложении в форме скалярных произведений (рис. 1) свидетельствует о том, что при изменении состояний КЭСК повторное разложение следует проводить не для всей матрицы $[K]$, а только для ее части, связанной с узлами КЭСК. В приведенном примере с использованием классического АМС при повторном разложении пришлось бы вычислить величины семи элементов треугольного множителя (8), а с использованием модифицированного АМС – только для трех (10).

Представленный подход был реализован в комплексе проблемно-ориентированных программ [9], разработанном для анализа динамического поведения многокаскадных сборных роторов турбомашин.

Численный эксперимент и анализ эффективности вычислительного алгоритма на конечноэлементных моделях сборных роторных систем. Для тестирования эффективности модифицированного АМС была разработана КЭ модель сборного ротора, эскиз которой показан на рис. 7а. Моделирование основных деталей ротора реализовано с применением КЭ типа «изопараметрический гексаэдр», контактных сопряжений – с применением КЭСК, моделирование крепежа деталей – с применением КЭСК, балочных КЭ Тимошенко и КЭ типа «невесомый стержень», а моделирование упругих опор – с применением упруго-демпферных КЭ [8]. КЭ модель представлена на рис. 7б.

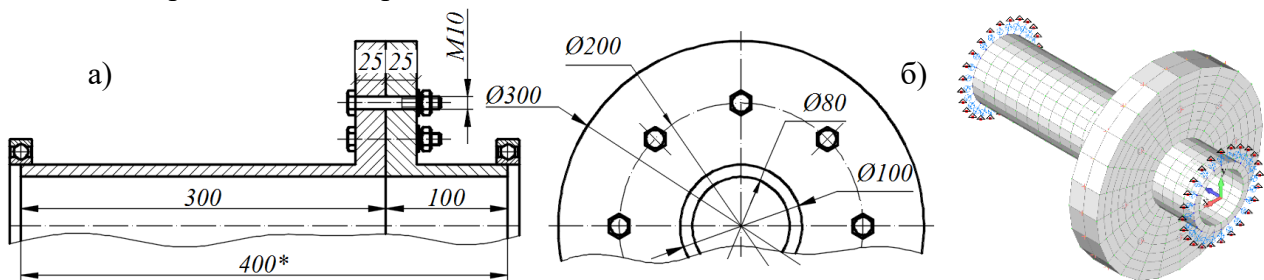


Рис. 7. Эскиз и конечноэлементная модель тестового сборного ротора

Всего представленная КЭ модель включает в себя: 1888 узлов (48 из которых закреплены по трем поступательным степеням свободы) и 1200 КЭ: 864 восьмиузловых КЭ типа «изопараметрический гексаэдр», 48 упруго-демпферных КЭ, 248 КЭСК, 24 балочных КЭ Тимошенко, и 16 КЭ типа «невесомый стержень». Доля узлов КЭСК составляет 27% от общего числа незакрепленных узлов модели. С учетом «вычеркивания» закрепленных степеней свободы матрица жесткости имеет размерность $n = 5520$.

На рис. 8а показан портрет исходной матрицы жесткости (главная диагональ и поддиагональные элементы), упорядоченный с использованием классического АМС, содержащий 147.912 элементов, а на рис. 8б – портрет треугольного множителя Холецкого, содержащий 933.585 элементов. Построение портретов матриц реализовано с применением языка программирования Python [16] и библиотеки «Pillow» для работы с изображениями; каждый пиксель черного цвета обозначает наличие вещественного ненулевого числа в матрице. Всего представленная матрица 5520×5520 на главной диагонали и под ней содержит 15.237.960 элементов.

При разложении Холецкого было подсчитано количество арифметических операций: 5.520 извлечений квадратного корня, 130.686.523 умножений, 130.686.523 сложений и вычитаний, 922.546 делений – всего 262.301.112 операции.

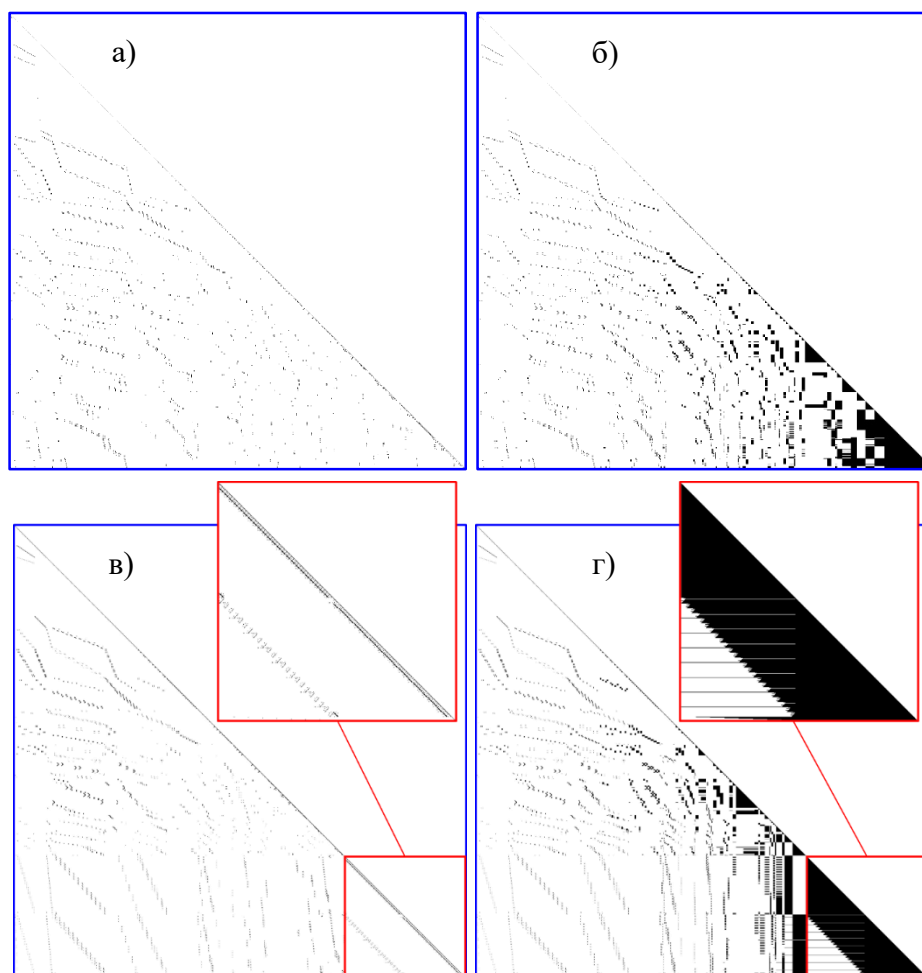


Рис. 8. Портреты матриц с использованием классического и модифицированного АМС

С использованием дополнительных тестовых примеров было установлено время простых арифметических операций, приведенное в секундах для 10^{10} операций в табл. 1.

Таблица 1. Время выполнения 10^{10} арифметических операций (FORTRAN), сек

Операция	Сложение	Вычитание	Умножение	Деление	Извлечение квадратного корня
Опыт 1	3,50	3,54	3,50	9,06	12,74
Опыт 2	3,45	3,48	3,45	9,03	12,81
Опыт 3	3,47	3,51	3,65	9,04	12,75
Опыт 4	3,47	3,50	3,44	9,05	12,75
Опыт 5	3,44	3,50	3,46	9,04	12,75
Среднее значение	3,49			9,04	12,76

Таким образом, условное время выполнения одной итерации разложения при использовании классического АМС (без учета других операций, например, присваивания, поиска элемента массива и т.п.) составляет 92,078 миллисекунды.

На рис. 8в показан портрет исходной матрицы жесткости, упорядоченный с использованием **модифицированного** АМС, содержащий 147912 элементов, а на рис. 8г – портрет треугольного множителя Холецкого, содержащий 1577346 элементов.

Количество арифметических операций при полном разложении Холецкого составляет: 5520 извлечений квадратного корня, 430524817 умножений, 430524817 сложений и вычитаний, 1566307 делений – всего 862621461 операции.

Условное время выполнения первой итерации (полного) разложения при использовании модифицированного АМС без учета других операций составляет 301,99 миллисекунды. Таким образом, время, затраченное при разложении Холецкого на первой итерации с упорядочением

структуры матрицы модифицированным АМС, в 3,28 раз превышает время, затраченное при разложении с упорядочением классического АМС.

Однако количество арифметических итераций и время повторных разложений оказываются несоизмеримо меньшими времени первой итерации. На рис. 8в и рис. 8г соответственно дополнительно выделены красным цветом портреты повторно раскладываемой части исходной матрицы и треугольного множителя Холецкого.

При этом в исходной матрице на главной диагонали и под ней содержится 21840 ненулевых элементов, а в треугольном множителе – 832416 ненулевых элементов. Количество арифметических операций при повторном (неполном) разложении Холецкого составляет: 1488 извлечений квадратного корня, 1022 умножений, 1022 сложений и вычитаний, 829441 делений – всего 832973 операции. Условное время выполнения одной итерации неполного разложения составляет 0,753 миллисекунды.

Таким образом, эффективность применения модифицированного АМС по сравнению с классическим АМС проявляется при решении задачи за четыре и более итераций (рис. 9).

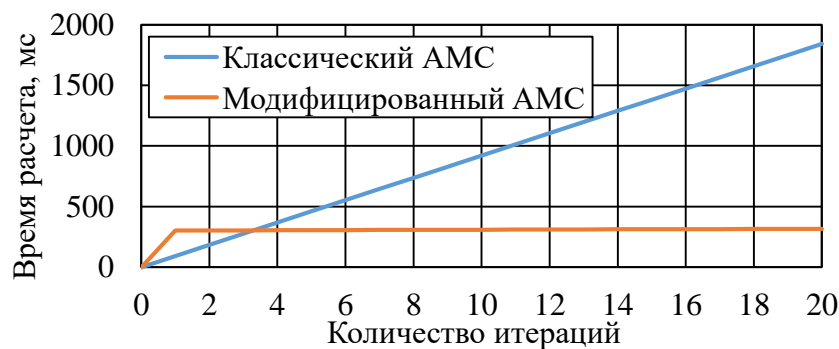


Рис. 9. Зависимость времени разложения от количества итераций при использовании разных методов упорядочения структуры исходной матрицы

При анализе динамического поведения сборных роторов турбомашин сходимость решения контактной задачи, оцениваемая по относительному количеству КЭСК с неизменным состоянием по сравнению с предыдущей итерацией, достигается за 6-10 итераций [14]. Кроме того, в некоторых случаях возникают ситуации неустойчивого поведения КЭСК, при котором несколько таких конечных элементов изменяют свое состояние от итерации к итерации, не позволяя завершить решение. В таких случаях количество итераций решения контактной задачи ограничено (двадцатью) и может быть изменено на этапе настройки решения. Таким образом, использование модифицированного АМС для упорядочения дает ощутимую экономию времени расчета (в представленном примере – в 1,81 раз при 6 итерациях, в 2,98 раза при 10 итерациях и в 5,82 раза при 20 итерациях).

Также с использованием тестовой КЭ модели было исследовано изменение времени выполнения полного и частичного разложения Холецкого при изменении доли общего числа КЭСК, но с сохранением общего количества узлов и размерности матриц. Для этого независимо от физической стороны задачи было проведено разложение для КЭ модели с 32, 56...392 КЭСК с шагом 24 элемента (соответствующему количеству узлов в окружном направлении ротора) и оценено количество итераций, на которых одно полное и несколько неполных разложений при применении модифицированного АМС дают выигрыш во времени по сравнению с несколькими разложениями при применении классического АМС (рис. 10). Оценка количества итераций определялась на основе неравенства:

$$T + (N - 1)t < N\tau \quad (11)$$

где T и t – соответственно время одного полного и неполного разложения Холецкого с применением модифицированного АМС; N – количество итераций; τ – время одного полного разложения Холецкого с применением классического АМС.

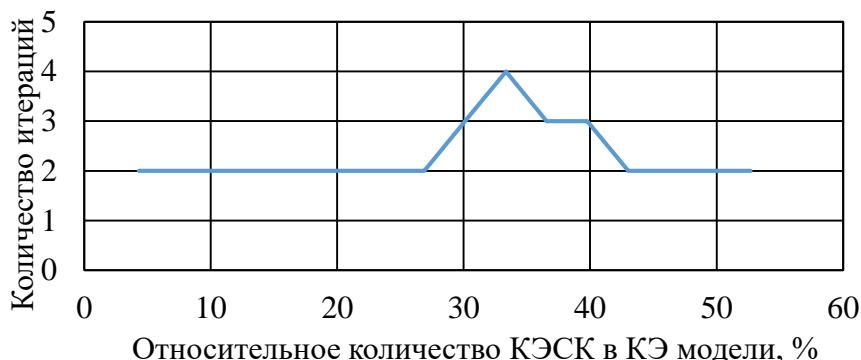


Рис. 10. Количество итераций решения контактной задачи, при котором упорядочение модифицированным АМС выгоднее упорядочения классическим АМС в зависимости от относительного количества КЭСК в модели

Исходя из полученных результатов можно сделать вывод, что эффективность упорядочения структуры матрицы жесткости модифицированным АМС может проявляться даже при большом относительном количестве КЭСК в модели. Однако на практике могут встречаться случаи, в которых модифицированный АМС приводит к большим затратам времени, вызванным вмешательством в работу основного АМС и неоптимальным (с точки зрения данного алгоритма) упорядочением.

Представленный в работе модифицированный АМС был применен для упорядочения структуры матрицы жесткости при анализе динамического поведения многокаскадной сборной роторной системы реального авиационного ГТД (рис. 11).

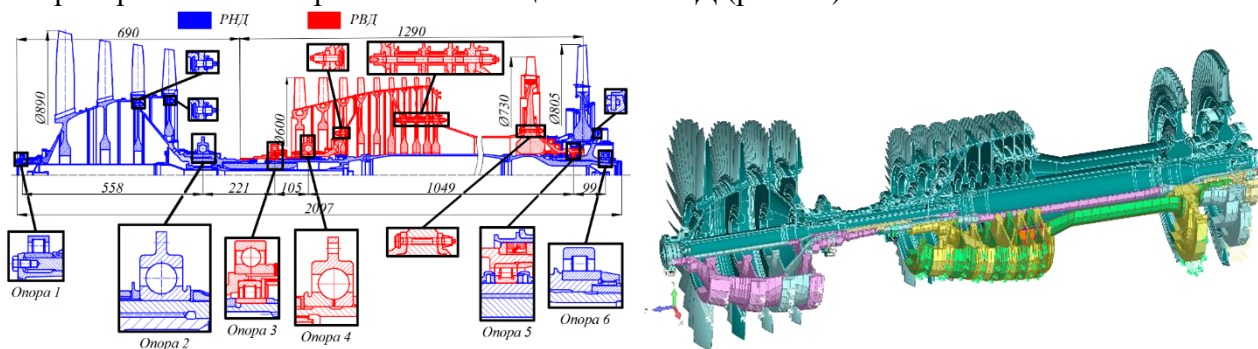


Рис. 11. Эскиз и КЭ модель, совмещенная с твердотельной моделью с вырезом одной четверти многокаскадной сборной роторной системы реального авиационного ГТД

Всего представленная КЭ модель включает в себя: 46596 узлов (104 из которых закреплено по трем поступательным степеням свободы), 30488 КЭ: 22568 восьмиузловых КЭ типа «изопараметрический гексаэдр», 104 упруго-демпферных КЭ, 3900 КЭСК, 1200 балочных КЭ Тимошенко, 2340 КЭ типа «невесомый стержень» и 376 КЭ типа «сосредоточенная масса». Доля узлов КЭСК составляет 25,6% от общего числа незакрепленных узлов модели. С учетом закрепленных степеней свободы матрица жесткости имеет размерность $n = 139476$.

На рис. 12а показан портрет исходной матрицы жесткости (главная диагональ и поддиагональные элементы), упорядоченный с использованием **классического** АМС, содержащий 3791880 элементов, а на рис. 12б – портрет треугольного множителя Холецкого, содержащий 31239756 элементов.

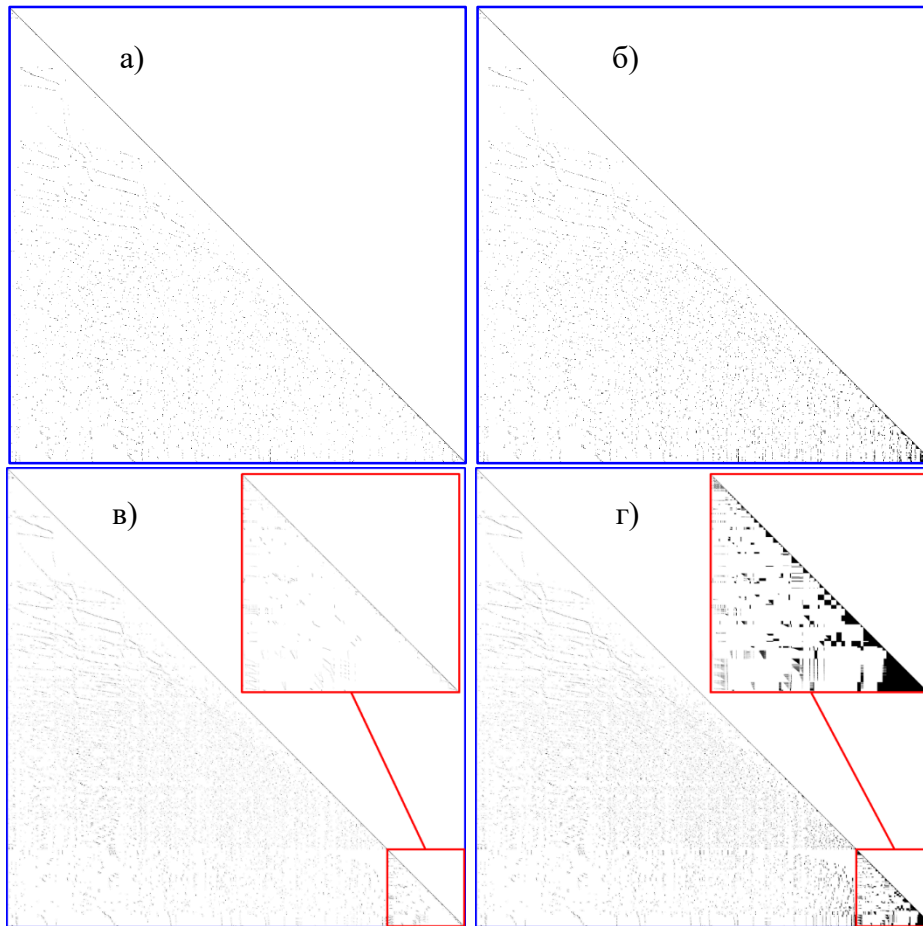


Рис. 12. Портреты матриц с использованием классического и модифицированного АМС для КЭ модели роторной системы реального ГТД

На рис. 12в показан портрет исходной матрицы жесткости, упорядоченный с использованием **модифицированного АМС**, содержащий 3791880 элементов, а на рис. 12г – портрет треугольного множителя Холецкого, содержащий 52541100 элементов.

Условное время выполнения полного разложения на первой итерации при использовании модифицированного АМС составляет 23,6151 с., время неполного разложения на последующих итерациях – 0,025014 с., а время полного разложения на каждой итерации при упорядочении классическим АМС составляет 6,4247 с. Таким образом, время, затраченное на разложение Холецкого в первой итерации при использовании упорядочения модифицированным АМС, в 3,68 раз, превышает время, затраченное при упорядочении классическим АМС. Эффективность применения модифицированного АМС, по сравнению с классическим АМС, проявляется при решении задачи за четыре и более итераций (рис. 13).

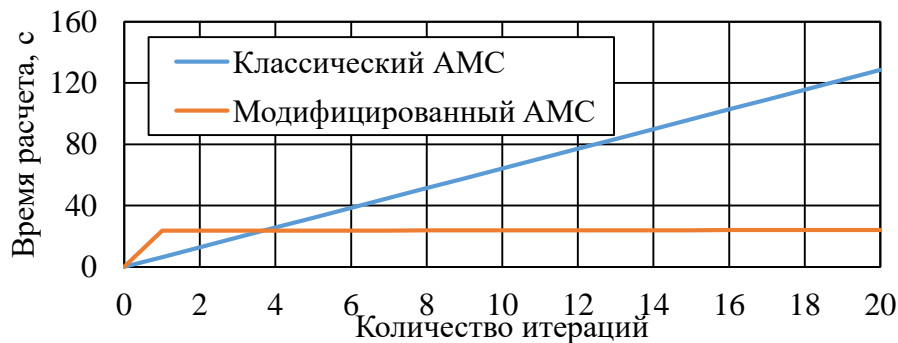


Рис. 13. Зависимость времени разложения от количества итераций при использовании разных методов упорядочения структуры исходной матрицы для КЭ модели роторной системы реального ГТД

Заключение. При сравнении результатов, полученных в работе, видно, что применение классического АМС упорядочивает структуру матрицы коэффициентов СЛАУ таким образом, что существенно позволяет уменьшить ее заполнение при разложении Холецкого, и, как следствие, уменьшить время разложения, что повышает эффективность вычислительного метода. Однако решение контактных задач итерационным способом можно проводить более эффективно (с меньшими вычислительными затратами) за счет проведения неполного разложения Холецкого на повторных итерациях при условии, что количество узлов, связанных с КЭС, существенно меньше общего количества узлов КЭ модели; упорядочение структуры матрицы коэффициентов СЛАУ в таком случае выполняется модифицированным АМС. Таким образом, при решении контактной задачи за несколько итераций применение модифицированного АМС в совокупности с частичным разложением может существенно повышать эффективность вычислительного алгоритма. Поскольку при этом структура матрицы упорядочивается не самым эффективным с точки зрения классического АМС способом, может возникать чрезмерное заполнение. Поэтому в программной реализации решения [9] проводятся два символических разложения Холецкого (с применением классического и модифицированного АМС) и оценивается меньшее время решения задачи. Однако даже для КЭ модели роторной системы реального ГТД с большим количеством элементов множителя Холецкого и арифметических операций для его получения более высокая эффективность достигается при применении модифицированного АМС с учетом решения контактной задачи за 6-10 итераций [14].

Список источников

1. Bathe K.J. Finite element procedures. Upper Saddle River, New Jersey, Prentice hall, 1996, 1038 p.
2. Гантмахер Ф.Р. Теория матриц / Ф.Р. Гантмахер. – Изд. 5-е. – Москва : Физматлит, 2004. – 559 с.
3. Расчеты машиностроительных конструкций методом конечных элементов: Справочник / В.И. Мяченков, В.П. Мальцев, В.П. Майборода и др.; под общ. ред. М.И. Мяченкова. – М.: Машиностроение, 1989. – 520 с.
4. Фиалко С.Ю. Прямые методы решения систем линейных уравнений в современных МКЭ-комплексах / С.Ю. Фиалко. – М.: Издательство СКАД СОФТ, Издательство Ассоциации строительных вузов (АСВ), 2009. – 160 с.
5. Джордж А. Численное решение больших разреженных систем уравнений: Пер. с англ. / А. Джордж, Дж. Лю. – М.: Мир, 1984. – 333с., ил.
6. Исследование структурных свойств алгоритма разложения Холецкого: от давно известных фактов до новых выводов / А.В. Фролов, В.В. Воеводин, И.Н. Коньшин, А.М. Теплов // Вестник Уфимского государственного авиационного технического университета, 2015. – Т. 19. – № 4(70). – С. 149-162.
7. Пирова А.Ю. Параллельные алгоритмы и программные средства для уменьшения заполнения фактора симметричных разреженных матриц: диссертация на соискание ученой степени кандидата технических наук / Пирова Анна Юрьевна, 2022. – 143 с.
8. Дудаев М.А. Математическое моделирование динамического поведения двухроторных систем турбомашин с межвальным подшипником / М.А. Дудаев, А.А. Пыхалов // Труды МАИ, 2024. – № 2.
9. Свидетельство о государственной регистрации программы для ЭВМ № 2019617798. Конечноэлементный решатель задачи роторной динамики одно- и двухвальных турбомашин с контактным взаимодействием деталей и межвальными связями: заявка № 2019616638 от 11.06.2019 РФ / М.А. Дудаев (РФ) – Зарегистрировано в Реестре программ для ЭВМ 20.06.2019 г. (РФ).
10. Бартенев О.В. Современный Фортран. – 3-е изд., доп. и перераб. / О.В. Бартенев. – М.: ДИАЛОГ-МИФИ, 2000. – 449 с.
11. Белоусов А.И. Дискретная математика / А.И. Белоусов, С.Б. Ткачев. – 7-е издание, исправленное. – Москва: Издательство МГТУ им. Н. Э. Баумана, 2021. – 704 с. – ISBN 5-7038-1769-2
12. Герб А.Р. Применение теории графов в алгебраических многосеточных методах для решения разреженных СЛАУ / А.Р. Герб, Г.А. Омарова // Проблемы информатики, 2022. – № 3(56). – С. 77-89.
13. Дудаев М.А. Контактная задача в анализе динамического поведения сборных роторов турбомашин / М.А. Дудаев, А.А. Пыхалов // Вестник НГТУ, 2015. – № 3. – С. 113-129.
14. Милов А.Е. Контактная задача динамики сборных роторов турбомашин : диссертация на соискание ученой степени кандидата технических наук / Милов Александр Евгеньевич, 2007. – 174 с.

15. Пыхалов А.А. Контактная задача статического и динамического анализа сборных роторов турбомашин: диссертация на соискание ученой степени доктора технических наук / А.А. Пыхалов Анатолий Александрович, 2006. – 405 с.

16. Бизли Д. Python: Подробный справочник / Д. Бизли. – 4-е изд. – Санкт-Петербург: Символ, 2010. – 858 с.

Дудаев Михаил Алексеевич. Старший преподаватель Иркутского государственного университета путей сообщения. Основные направления исследований автора: механика деформируемого твердого тела, динамика, метод конечных элементов. AuthorID: 688571, SPIN: 7575-9125, dudaev_ma@mail.ru, 664074, г. Иркутск, ул. Чернышевского, 15.

UDC 519.673

DOI:10.25729/ESI.2024.35.3.012

The minimum degree algorithm modification for improving the computational efficiency in the iterative solution of the contact problem of elasticity theory using the finite element method

Mikhail A. Dudaev

Irkutsk state transport university, Russia, Irkutsk, dudaev_ma@irgups.ru

Abstract. The article proposes a modification to the classical minimum degree algorithm, designed to optimize the structure of the global coefficient matrix in the system of linear algebraic equations used in the finite element method. The aim is to improve the efficiency of the computational algorithm when solving contact problems in elasticity theory, specifically when modeling the connection of parts using special contact elements that determine the iterative solution with a gradual refinement of the state. The proposed modification is implemented in conjunction with the direct Cholesky method for solving systems of linear algebraic equations. The computational efficiency of this method is demonstrated to the greatest extent when applied to systems with a symmetric, positive-definite coefficient matrix, which often has a sparse structure when using the finite element method. A significant problem in the Cholesky decomposition method is the filling of the "triangular multiplier" relative to the original matrix during the decomposition process. The reduction of this filling is typically achieved using well-known heuristic algorithms, among which the minimum degree algorithm has become widely used. The modification of the algorithm proposed in this work, by symmetric permutation of rows and columns of the initial global matrix, reorders the structure of this matrix in a way that, during the iterative solution of contact problems of elasticity theory, allows for partial Cholesky factorization to be performed in repeated iterations, thereby saving calculation time compared to performing a full factorization. The paper demonstrates that overall the matrix structure ordered by the modified algorithm is less preferable than that produced by the classical minimum-degree algorithm, since it is undergoing more filling. However, the efficiency of the computational method is achieved through repeated iterations that gradually specified the state of the contact finite elements of the structures coupling. Using a test finite element model of a rotor, as well as a model of the real aviation gas turbine engine, portraits of the initial matrices and their triangular multipliers are determined and presented, estimates of the number of arithmetic operations and the estimated time required for their execution are calculated using both a classical and a modified minimum degree algorithm.

Keywords: finite element, finite element method, contact problem, Cholesky method, minimum degree algorithm, scalar products, filling, symmetric permutation

References

1. Bathe K.J. Finite element procedures. Upper Saddle River, New Jersey, Prentice hall, 1996, 1038 p.
2. Gantmaher F.R. Teorija matric [Matrix theory]. Izd. 5-e., Moskva, Fizmatlit [Ed. 5th., Moscow, Fizmatlit], 2004, 559 p.
3. Mjachenkov V.I., Mal'cev V.P., Majboroda V. P. et al. Raschety mashinostroitel'nyh konstrukcij metodom konechnykh jelementov: spravochnik [Calculations of machine-building structures by the finite element method: Handbook]. M., Mashinostroenie [Mechanical engineering], 1989, 520 p.

4. Fialko, S. Ju. Prjamye metody reshenija sistem linejnyh uravnenij v sovremennyh MKJe-kompleksah [Direct methods for solving systems of linear equations in modern FEM complexes]. M., Izdatel'stvo SKAD SOFT, Izdatel'stvo Associacii stroitel'nyh vuzov (ASV) [M., Publishing house SKAD SOFT, Publishing house of the Association of construction universities (ACS)], 2009, 160 p.
5. Dzhordzh A. Chislennoe reshenie bol'shih razrezhennyh sistem uravnenij: Per. s angl. [Numerical solution of large sparse systems of equations: transl. from English]. M., Mir, 1984, 333p.
6. Frolov A.V., Voevodin V.V., Kon'shin I.N., Teplov A.M. Issledovanie strukturnyh svojstv algoritma razlozhenija Holec'kogo: ot davno izvestnyh faktov do novyh vyvodov [Study of structural properties of cholesky decomposition: from well-known facts to new conclusions] Vestnik Ufim'skogo gosudarstvennogo aviacionnogo tehničeskogo universiteta [Bulletin of the Ufa state aviation technical university], 2015, vol. 19, no. 4(70), pp. 149-162.
7. Pirova A.Ju. Parallel'nye algoritmy i programmnye sredstva dlja umen'shenija zapolnenija faktora simmetričnyh razrezhennyh matric: dissertacija na soiskanie uchenoj stepeni kandidata tehničeskikh nauk [Parallel algorithms and software tools for reducing the filling factor of symmetric sparse matrices : dissertation for the degree of Candidate of technical sciences], 2022. – 143 p.
8. Dudaev M.A., Pyhalov A.A. Matematičeskoe modelirovanie dinamičeskogo povedenija dvuhrotornyh sistem turbomashin s mezhval'nym podshipnikom [The mathematical modeling of the dynamic behavior of twin-rotor turbomachine systems with an inter shaft bearing]. Trudy MAI, 2024, no. 2.
9. Svidetel'stvo o gosudarstvennoi registratsii programmy dlja EVM №2019617798. Konečnoelementnyi reshatel' zadachi rotornoj dinamiki odno- i dvukhval'nykh turbomashin s kontaktym vzaimodejstviem detalei i mezhval'nymi svyaziami: zaiavka №2019616638 ot 11.06.2019 RF [Certificate of state registration of the computer program No. 2019617798. Finite element solver of the problem of rotor dynamics of single- and two-shaft turbomachines with contact interaction of parts and inter-shaft connections: application No. 2019616638 dated 06/11/2019 of the Russian Federation]. Dudaev M.A. Zaregistrovano v reestre programm dlja EVM 20.06.2019 g. (RF).
10. Barten'ev O. V. Sovremennyy Fortran [Modern Fortran]. M., DIALOG-MIFI, 2000, 449 p.
11. Belousov A.I., Tkachev S.B. Diskretnaja matematika [Discrete mathematics]. 7-e izdanie, ispravlennoe. Moskva, Izdatel'stvo MGTU im. N. Je. Baumana [7th edition, revised. Moscow, Publishing house of Bauman Moscow state technical university], 2021, 704 p., ISBN 5-7038-1769-2.
12. Gerb A.R., Omarova G.A. Primenenie teorii grafov v algebraičeskikh mnogosetočnyh metodah dlja reshenija razrezhennyh SLAU [Application of graph theory in algebraic multigrid methods for solving sparse SLAEs]. Problemy informatiki [Problems of informatics], 2022, no. 3(56), pp. 77-89.
13. Dudaev M.A., Pyhalov A.A. Kontaktnaja zadacha v analize dinamičeskogo povedenija sbornyh rotorov turbomashin [The contact problem in the analysis of the dynamic behavior of modular turbomachine rotors]. Vestnik NGTU, 2015, no. 3. pp. 113-129.
14. Milov A.E. Kontaktnaja zadacha dinamiki sbornyh rotorov turbomashin : dissertacija na soiskanie uchenoj stepeni kandidata tehničeskikh nauk [The contact problem of the dynamics of prefabricated rotors of turbomachines : dissertation for the degree of candidate of technical sciences]. Milov Aleksandr Evgen'evich, 2007, 174 p.
15. Pyhalov A.A. Kontaktnaja zadacha statičeskogo i dinamičeskogo analiza sbornyh rotorov turbomashin : dissertacija na soiskanie uchenoj stepeni doktora tehničeskikh nauk [The contact problem of static and dynamic analysis of combined rotors of turbomachines : dissertation for the degree of doctor of technical sciences]. Pyhalov Anatolij Aleksandrovich, 2006, 405 p.
16. Bizli D. Python: Podrobnyy spravochnik [Python: A detailed reference guide]. Sankt-Peterburg, Simvol [Symbol], 2010, 858 p.

Dudaev Mikhail Alekseevich. Assistant professor Irkutsk state transport university, The main areas of research of the author: mechanics of a deformable solid body, dynamics, finite element method. AuthorID: 688571, SPIN: 7575-9125, dudaev_ma@mail.ru, 664074, Irkutsk, Chernyshevsky str., 15.

Статья поступила в редакцию 13.05.2024; одобрена после рецензирования 01.10.2024; принята к публикации 10.10.2024.

The article was submitted 05/13/2024; approved after reviewing 10/01/2024; accepted for publication 10/10/2024.

УДК 004.942

DOI:10.25729/ESI.2024.35.3.013

Планирование поставок продукции на год с учетом минимизации затрат на перевозку для бесперебойного снабжения потребителей

Маланова Татьяна Валерьевна, Бахвалов Сергей Владимирович,
Янхаев Данил Вадимович, Аршинский Вадим Леонидович

Иркутский национальный исследовательский технический университет,
Россия, Иркутск, *malanova_tanya@mail.ru*

Аннотация. В статье рассматривается задача планирования перевозки груза от поставщиков к потребителям на год, поделенный на несколько периодов, в течение которых нужно вывезти определенную часть товаров. Товар вывозится за несколько рейсов. Для решения этой задачи предлагается использовать две модели транспортной задачи линейного программирования, в каждую из которых добавлено ограничение по времени. В одной из них переменными являются количества рейсов, а в другой – количества перевезенного груза от каждого потребителя к каждому поставщику. Для корректного применения указанных моделей в каждом периоде предлагается подготовить входные данные путем проведения предварительного этапа, обеспечивающего выполнение условия баланса транспортной задачи. Продемонстрирован пример.

Ключевые слова: задача линейного программирования, транспортная задача, математическое моделирование, исследование операций, управление перевозками

Цитирование: Маланова Т.В. Планирование поставок продукции на год с учетом минимизации затрат на перевозку для бесперебойного снабжения потребителей / Т.В. Маланова, С.В. Бахвалов, Д.В. Янхаев, В.Л. Аршинский // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 148-156. – DOI: 10.25729/ESI.2024.35.3.013.

Введение. Для крупных предприятий и государственных учреждений характерно планирование перевозок товаров от поставщиков до потребителей в течение заданного времени, например, годовое планирование. Если рассматривать товары первой необходимости или сырье для производства, то поставка товаров должна обеспечивать бесперебойное их потребление. В то же время некоторые пункты потребления могут быть доступны не всегда. Так, например, перевозки невозможны в периоды, когда не работают паромные или ледовые переправы через реки. Для северных районов часто перевозка может быть осуществлена только по зимней дороге или только по рекам летом. Поэтому необходимо при планировании учесть, что в доступное для перевозки время товар должен поставляться с излишком, чтобы во время отсутствия дороги сырье или необходимые товары были в наличии.

Потребности для потребителей и наличие товаров у поставщиков должны быть известными величинами. Так, например, потребность в товарах первой необходимости для населения можно вычислить, исходя из его численности и суточных нормативов потребления; потребность в сырье для производства определяется по объемам и соответствующим нормам производства. Перевозка груза может осуществляться в несколько рейсов, что приводит к необходимости учитывать время доставки, если установлены сроки перевозки товаров.

Затраты на перевозку грузов от заданного поставщика (склада) до пункта потребления определяются протяженностью дорог, их качеством и видом транспорта. В данной статье учитывается условие, что при транспортировке товаров большими объемами часто оплата производится не по объему перевезенного груза, а по количеству произведенных рейсов. Учитывая ограниченность финансовых ресурсов, разработанный план должен обеспечить минимальные (или близкие к минимуму) суммарные затраты на перевозку товаров в течение определенного времени.

Если поставщиков и потребителей достаточно много, то такую задачу вручную качественно решить невозможно, поэтому разработка программного обеспечения для планирования поставок товаров является актуальной задачей.

На сегодняшний день существует программное обеспечение, использующее методы математического моделирования в управлении поставками, такие, как Forecast NOW, ConsID, GoodsForecast.Replenishment, GoodsForecast.Distribution, TMS Логистика. Все они направлены на оптимизацию задач бизнеса и минимизацию затрат, но не учитывают специфику задач планирования для госучреждений и госпредприятий – необходимо бесперебойное обеспечение всех потребителей, даже в тех случаях, когда это может быть экономически не оправдано, что не позволяет применить их для решения задачи в заданном контексте.

В литературе рассматриваются транспортные задачи с различными их модификациями: в классической постановке [1, 2, 3]; с дополнительными ограничениями [4, 5]; как многокритериальные оптимизационные задачи [6, 7]. Также рассматриваются задачи маршрутизации транспорта [8, 9, 10, 11], но они не решают задачу планирования на период (например, на год) с учетом доступности/недоступности потребителей в определенные сезоны. Таким образом, в данной статье предлагается алгоритм формирования плана поставки товаров от поставщиков до потребителей (с учетом временной недоступности некоторых из них) в течение года или в более общей постановке – в течение заданного периода таким образом, чтобы потребление было непрерывным.

1. Математическая модель. Для удобства изложения примем период планирования за год. В течение года поставка может осуществляться партиями (ежемесячно, еженедельно, ежедневно). Поэтому весь год необходимо разделить на l промежутков для поставки продукции (например, на 12 месяцев). Данные промежутки определяются, исходя из конкретной задачи, в зависимости от срока годности товаров, от периода отчетностей на предприятии и т.п.

В каждом промежутке решается транспортная задача с ограничением по времени. Математическая модель, переменными которой являются количества рейсов от поставщиков к потребителям, имеет вид:

$$F(r) = \sum_{i=1}^m \sum_{j=1}^n (c_{ij} r_{ij}) \rightarrow \min \quad (1)$$

при ограничениях

$$\begin{cases} \sum_{j=1}^n (g_{ij} r_{ij} - z_{ij}) = a_i, & i = \overline{1, m} \\ \sum_{i=1}^m (g_{ij} r_{ij} - z_{ij}) = b_j, & j = \overline{1, n} \\ \sum_{j=1}^n r_{ij} t_{ij} \leq s, & i = \overline{1, m} \\ g_{ij} r_{ij} - z_{ij} > 0, & i = \overline{1, m}, j = \overline{1, n} \\ r_{ij} \geq 0, & i = \overline{1, m}, j = \overline{1, n} \\ z_{ij} > 0, & i = \overline{1, m}, j = \overline{1, n}. \end{cases} \quad (2)$$

Здесь считаем, что c_{ij} – транспортные расходы из i -го пункта в j -й пункт; r_{ij} – количество рейсов, произведенных в течение текущего промежутка; t_{ij} – время, за которое выполняется один рейс; g_{ij} – грузоподъемность транспорта;

$$z_{ij} = g_{ij} r_{ij} - x_{ij} \quad (3)$$

где x_{ij} – объем перевозимого груза из i -го пункта в j -й пункт, т.е. это объем груза, которого не хватает до полной загрузки транспорта; s – длина промежутка времени; b_j – объем заявок; a_i – объем запасов; m – число складов; n – число пунктов потребления.

Математическая модель (1), (2) представляет собой модель транспортной задачи (Т-задачи), зависящей от количества рейсов, с ограничением по времени. Все координаты решения данной задачи должны принадлежать множеству целых неотрицательных чисел. В связи с

этим количество вычислений может быть гораздо больше, чем в обычных задачах линейного программирования без условия на целочисленность. Так как количество переменных здесь $2 \times m \times n$, то при достаточно больших значениях m, n время решения задачи может существенно увеличиться.

Для ускорения расчетов сократим количество переменных модели (1), (2). Согласно равенству (3), перейдем от переменных r_{ij} и z_{ij} к x_{ij} , тем самым незначительно увеличив погрешность расчета (погрешность данной процедуры составляет z_{ij}). После преобразований целевая функция примет вид:

$$\sum_{i=1}^m \sum_{j=1}^n \frac{c_{ij}}{g_{ij}} x_{ij} \rightarrow \min, \quad (4)$$

а ограничения:

$$\begin{cases} \sum_{j=1}^n x_{ij} = a_i, i = \overline{1, m}, \\ \sum_{i=1}^m x_{ij} = b_j, j = \overline{1, n}, \\ \sum_{j=1}^n \frac{t_{ij}}{g_{ij}} x_{ij} \leq s, i = \overline{1, m}, \\ x_{ij} > 0, i = \overline{1, m}, j = \overline{1, n}. \end{cases} \quad (5)$$

Заметим, что при округлении на прошлых этапах условие баланса T-задачи частично не соблюдается. Возникает случай, при котором сумма запасов незначительно превышает сумму запросов, для учета данного случая преобразуем ограничения (4) и получим:

$$\begin{cases} \sum_{j=1}^n x_{ij} \leq a_i, i = \overline{1, m}, \\ \sum_{i=1}^m x_{ij} = b_j, j = \overline{1, n}, \\ \sum_{j=1}^n \frac{t_{ij}}{g_{ij}} x_{ij} \leq s, i = \overline{1, m}, \\ x_{ij} > 0, i = \overline{1, m}, j = \overline{1, n}. \end{cases} \quad (6)$$

Задача (4), (6) содержит в два раза меньше переменных, чем первоначальная модель, что поможет ускорить процесс вычисления. Кроме того, если количество перевозимого груза считается в единицах массы, то такая задача является классической задачей линейного программирования, в которой переменные принадлежат множеству действительных чисел, и это еще более ускорит процесс вычисления. Но, так как в настоящей работе мы считаем, что затраты и время перевозки вычисляются по количеству рейсов, то модель (1), (2) дает нам более точное решение, чем модель (4), (6).

Таким образом, в данной статье для решения поставленной задачи предлагается использовать одну из двух моделей в зависимости от необходимой точности времени на вычисления.

2. Алгоритм. Как известно, чтобы T-задача имела решение, необходимо, чтобы выполнялось условие баланса – общий объем заявок должен быть равен суммарному объему товаров на складах. Поэтому для годового плана, а также для всех периодов, необходимо выполнение условия баланса, то есть нужно привести исходные данные по наличию и заявкам к такому виду, чтобы T-задача решалась корректно в каждом периоде. Простым делением количества запасов и заявок на l периодов здесь невозможно обойтись, так как в определенных периодах некоторые потребители могут быть недоступными, то есть количество промежутков для поставки товаров этим потребителям меньше, чем l .

На основе выше сказанного алгоритм состоит из двух этапов: 1) подготовительный; 2) решение T-задачи с ограничением по времени.

2.1. Первый этап. На первом этапе предлагается равномерно распределить годовые запасы и запросы по периодам с учетом недоступности потребителей в определенные периоды, что с экономической точки зрения наиболее реально.

Пусть необходимо распределить годовой план на l периодов равномерно. Необходимо перераспределить товар, который предназначен для потребления в недоступные периоды, на доступные периоды, то есть нужно сделать так, чтобы суммарный объем запросов за все периоды каждого отдельного потребителя был равен годовому объему запросов этого потребителя; аналогичное условие и для поставщиков. При этом должно быть выполнено условие баланса: годовой запас товара у поставщиков должен быть равен общему годовому объему запросов.

Алгоритм распределения товаров включает в себя несколько шагов:

- 1) равномерно распределяем запросы каждого потребителя по периодам;
- 2) для каждого потребителя, у которых есть недоступные периоды (назовем их «закрытыми»), подсчитываем объем товара, который должны перебросить на каждый доступный период («открытый» период), предшествующий «закрытому», и производим перераспределение объема товара с каждого «закрытого» периода каждого такого потребителя на все предшествующие «открытые»;
- 3) равномерно распределяем запасы по периодам;
- 4) перебрасываем равномерно весь товар с «закрытых» периодов первого потребителя для всех поставщиков, то есть объем перебрасываемого товара в соответствующем «закрытом» периоде нужно разделить на количество поставщиков и количество предшествующих открытых периодов, затем полученный объем добавить к каждому поставщику во всех предшествующих «открытых» периодах.
- 5) далее повторяем 4 этап для каждого следующего потребителя, у которых были «закрытые» периоды.

Таким образом, после выполнения данного алгоритма мы получим в каждом периоде закрытую транспортную задачу.

При равномерном распределении ресурсов полученные значения не являются целочисленными, поэтому, если необходимо, то можно округлить каждое значение в большую сторону. В этом случае необходимо соблюдать условие: объем запросов отдельного периода не должен превышать объема запасов этого же периода. При нарушении данного условия можно добавлять к каждому поставщику единицу товара, до тех пор, пока условие баланса не будет выполнено (погрешность при данной процедуре, не превышает количества n – потребителей).

Пример. Рассмотрим простейший пример перераспределения товаров между поставщиками и потребителями по периодам.

На рисунке 1 представлен годовой план перевозки товаров из трех складов четырем потребителям.

	B1	B2	B3	B4	ΣA_i
A1					12000
A2					19600
A3					15200
ΣB_j	12800	11600	12000	10400	

Рис. 1. Годовой план перевозки товаров

Пусть поставка осуществляется за четыре периода, и известна доступность потребителей в каждом периоде:

- в первом периоде все потребители доступны;
- во втором периоде недоступен пункт B_4 ;
- в третьем периоде недоступны пункты B_2 и B_4 ;
- в четвертом периоде все потребители доступны.

Согласно алгоритму сначала распределяем запросы равномерно по всем периодам, рисунок 2(а). Затем перебрасываем товар из ячейки (3,2) во все ячейки столбца B_2 , соответствующим периодам, предшествующим третьему, в размере $2900/2$ периода = 1450. Аналогично поступаем с B_4 : перебрасываем товар со второго и третьего периодов в первый в размере $2600 \cdot 2$ периода = 5200. Ячейки, расположенные после недоступных периодов, не изменяем.

	B1	B2	B3	B4		B1	B2	B3	B4
1 период	3200	2900	3000	2600	1 период	3200	4350	3000	7800
2 период	3200	2900	3000	2600	2 период	3200	4350	3000	
3 период	3200	2900	3000	2600	3 период	3200		3000	
4 период	3200	2900	3000	2600	4 период	3200	2900	3000	2600

(а)

(б)

Рис. 2. Перераспределение заявок с учетом «закрытых» периодов у поставщиков B_2 и B_4

Далее перераспределяем товар по складам, рисунок 3. На этом шаге также сначала распределяем его равномерно по периодам, рисунок 3(а). Затем перебрасываем в ячейки первого и второго периодов весь товар из клеток третьего периода в размере $1450/3$ склада = 483,33. Аналогично перебрасываем товар из ячеек третьего и четвертого периодов в ячейки первого периода в размере $5200/3$ склада = 1766,66. В результате получили план хранения товаров у поставщиков на складах и подачи заявок на товар в каждом периоде.

	A1	A2	A3
1 период	3000	4900	3800
2 период	3000	4900	3800
3 период	3000	4900	3800
4 период	3000	4900	3800

(а)

	A1	A2	A3		A1	A2	A3
1 период	3483,33	5383,33	4283,33	1 период	5216,67	7116,67	6016,67
2 период	3483,33	5383,33	4283,33	2 период	2616,67	4516,67	3416,67
3 период	2033,33	3933,33	2833,33	3 период	1166,67	3066,67	1966,67
4 период	3000	4900	3800	4 период	3000	4900	3800

(б)

(в)

Рис. 3. Перераспределение запасов с учетом «закрытых» периодов у поставщиков B_2 и B_4

Следует отметить, что полученные данные несут рекомендательный характер для поставщиков и потребителей. Можно товары перераспределять и вручную по требованиям лица, принимающего решение, но нужно следить за тем, чтобы в каждом периоде и в целом за год выполнялось условие баланса для транспортной задачи.

После равномерного распределения ресурсов по периодам, необходимо убрать в соответствующих периодах недоступных потребителей. Таким образом, получаем следующие таблицы заявок от потребителей и наличия на складах в каждом периоде (рисунок 4) – исходные данные для решения Т-задач в каждом периоде.

	B1	B2	B3	B4	сум Ai		B1	B2	B3	сум Ai
A1					5216,67	A1				2616,67
A2					7116,67	A2				4516,67
A3					6016,67	A3				3416,67
сум Bj	3200	4350	3000	7800	18350	сум Bj	3200	4350	3000	10550

(а)

(б)

	B1	B3	сум Ai
A1			1166,67
A2			3066,67
A3			1966,67
сум Bj	3200	3000	6200

(в)

	B1	B2	B3	B4	сум Ai
A1					3000
A2					4900
A3					3800
сум Bj	3200	2900	3000	2600	11700

(г)

Рис. 4. Таблицы значений объемов запасов и заявок в каждом периоде

Предложенную схему распределения запасов и запросов можно использовать для годового планирования перевозок, начиная с любой даты.

На этом первый этап алгоритма заканчивается.

2.2. Второй этап. На втором этапе решается транспортная задача с ограничением по времени в каждом периоде [12]. В данном случае именно период и накладывает ограничение по времени. Как было уже сказано ранее, при использовании модели (1), (2) на этом этапе могут возникнуть проблемы, связанные с большой размерностью задачи. Так, задача 5x12 содержит 120 переменных, это приводит к тому, что время расчетов значительно увеличивается. Особенно это становится заметным, когда достаточно много периодов поставки в году, например, планирование осуществляется по месяцам, неделям.

Если не требуется высокой точности результата, то есть достаточно получить решение хотя бы близкое к оптимальному, то предлагается применить задачу (4), (6), которая решается быстрее, но дает квазиоптимальное решение. Ниже представлен код описания и решения поставленной задачи по второй модели на языке Python.

Листинг 1. Описание и решение модели (4), (6)

```
# Создаем модель
model = LpProblem("transport-problem-with-critical-time", LpMinimize)
# Инициализируем переменные решения:
X = {} # Будет хранить количества рейсов каждого маршрута
for a in A: # Заполняем словари переменными
    for b in B:
        var_name = "x_" + a + "_" + b # Название переменной
        X[a, b] = pulp.LpVariable(var_name, lowBound=0, cat="Integer") #
Инициализация переменной
# Задаем целевую функцию
lst_mult = [C[key] / G[key] * var for key, var in X.items()]
obj_expression = pulp.lpSum(lst_mult)
model.setObjective(obj_expression) # Добавили в модель

for a in A: # Создаем ограничение для каждого поставщика
    constr_name = f"{a}_constraint" # Название ограничения
    lst_vars = [X[a, b] for b in B]
    model += pulp.lpSum(lst_vars) <= expenses[a], constr_name # Все запасы
должны быть использованы

    constr_name = f"S{a}_constraint" # Название ограничения
    lst_vars = [X[a, b] * T[a, b] / G[a, b] for b in B]
    model += pulp.lpSum(lst_vars) <= s, constr_name # Затраты по времени каж-
дого поставщика должны не превышать s
```

```
for b in B: # Создаем ограничение для каждого потребителя
    constr_name = f"{b}_constraint" # Название ограничения
    lst_vars = [X[a, b] for a in A]
    model += pulp.lpSum(lst_vars) == expenses[b], constr_name # Все запросы
должны быть удовлетворены

# Решаем задачу оптимизации с помощью решателя
if timeCalculation != 0:
    model.solve(PULP_CBC_CMD(msg=1, maxSeconds=math.ceil(timeCalculation/periodCount)))
else:
    model.solve()
```

Заключение. В статье предложен алгоритм формирования годового плана поставок товаров поставщикам, которые могут быть недоступными в определенные периоды. Расход на перевозку товаров рассчитывается, исходя из количества рейсов, совершенных для перевозки грузов. Для расчета плана поставок в каждом периоде предложены две модели Т-задачи. В случае, когда организация строго ограничена в средствах, рекомендуется использовать более точную, исходя из данной постановки задачи, модель для вычисления оптимального опорного плана, переменные решения которой – количества рейсов. Если скорость вычисления наиболее значима, чем затраты на перевозку, то предлагается использовать для вычисления оптимального плана перевозок модель с сокращенным количеством переменных решения, обозначающих количество перевозимого груза. Чтобы в каждом периоде Т-задача решалась корректно, предложенный алгоритм предусматривает этап перераспределения товаров до вычисления опорных планов в каждом периоде.

Список источников

1. Болбаков Р.Г. Транспортная задача спроса и предложения / Р.Г. Болбаков, В.Я. Цветков // ИТНОУ: Информационные технологии в науке, образовании и управлении, 2018. – № 6 (10). – С. 83-88.
2. Зайченко Ю.П. Исследование операций: учеб. для вузов / Ю.П. Зайченко. – 3-е изд., перераб. и доп. – Киев: Выща шк., 1988. – 552 с.
3. Хэмди А. Таха Введение в исследование операций: в 2 кн. / Хэмди А. Таха; пер. с англ. В. Я. Алтаева и др. – М.: Мир, 1988. – 912 с.
4. Петрунин С.В. Решение транспортных задач ПС-методом при ограничениях на переменные // Научный вестник МГТУ ГА, 2014. – № 202.
5. Шапкин А.С. Математические методы и модели исследования операций: Учебник / А.С. Шапкин, В.А. Шапкин. – 7-е изд. – М.: Издательско-торговая корпорация «Дашков и Ко», 2019. – С. 398 с.
6. Золотарюк А.В. Математическая модель многокритериальной оптимизации транспортных перевозок. / А.В. Золотарюк // Инновационные технологии в науке и образовании, 2015. – № 1(1). – С. 317-320
7. Нечитайло Н.М. Задачи транспортного типа по критерию времени с учётом характеристик применяемых транспортных средств / Н.М. Нечитайло // Мир транспорта, 2021. – Т. 19. – № 3 (94). – С. 74-80.
8. Кубил В.Н. Обзор динамических задач маршрутизации транспорта / В.Н. Кубил, Ю.О. Чернышев // Программные продукты и системы, 2020. – Т. 33. – № 3. – С. 491-501.
9. Костюк Ю.Л. Приближенные алгоритмы решения сбалансированной задачи к коммивояжеров / Ю.Л. Костюк, М.С. Пожидаев // Вестник Томского государственного университета. Управление, вычислительная техника и информатика, 2008. – № 1(2). – С. 106–112.
10. Пожидаев М.С. Сбалансированная эвристика для решения задачи маршрутизации транспорта с учетом грузоподъемности / Ю.Л. Костюк, М.С. Пожидаев // Вестник ТГУ. УВТиИ. – 2010. – №3. – С. 65-72
11. Тюрин А.Ю. Модели транспортного обслуживания в цепях поставок пищевой промышленности / А.Ю. Тюрин // Вестн. КузГТУ, 2011. – № 4. – С.89-92.
12. Meindl B., Templ M. Analysis of commercial and free and open source solvers for linear optimization problems. Institut f. Statistik u. Wahrscheinlichkeitstheorie, Austria, 2012, 13 p.

Маланова Татьяна Валерьевна. К.т.н., Иркутский национальный исследовательский технический университет, доцент центра программной инженерии, AuthorID: 530825, ORCID: 0000-0002-1586-2530, malanova_tanya@mail.ru. 664074, г. Иркутск, ул. Лермонтова 83.

Бахвалов Сергей Владимирович. К.т.н., Иркутский национальный исследовательский технический университет, доцент центра программной инженерии, AuthorID: 169991, ORCID:0009-0009-83961926, bsv@istu.edu, 664074, г. Иркутск, ул. Лермонтова 83.

Янхаев Данил Вадимович. Магистрант, Иркутский национальный исследовательский технический университет, yanhaev2000@gmail.com, 664074, г. Иркутск, ул. Лермонтова 83.

Аршинский Вадим Леонидович. К.т.н., Иркутский национальный исследовательский технический университет, доцент центра программной инженерии, AuthorID: 520305. arshinskyv@istu.edu, ORCID: 0000-0001-7832-1582 664074, г. Иркутск, ул. Лермонтова 83.

UDC 004.942

DOI:10.25729/ESI.2024.35.3.013

Planning product supplies for the year, taking into account minimizing transportation costs

Tatyana V. Malanova, Sergey V. Bakhvalov, Danil V. Yanhaev, Vadim L. Arshinsky

Irkutsk national research state technical university,
Russia, Irkutsk, malanova_tanya@mail.ru

Abstract. The article examines the problem of planning the transportation of goods from suppliers to consumers for a year, divided into several periods, during which it is necessary to remove a certain part of the goods. The goods are exported over several flights. To solve this problem, it is proposed to use two models of the transport problem, each of which has a time constraint added. In one of them, the variables are the number of flights, and in the other, the amount of cargo transported from each consumer to each supplier. For the correct application of these models in each period, it is proposed to prepare input data by carrying out a preliminary stage that ensures the fulfillment of the balance condition of the transport task. An example is shown.

Keywords: linear programming problem, transportation problem, mathematical modeling, operations research, transportation management

References

1. Bolbakov R.G., Tsvetkov V. Ya., Bolbakov R.G. Transportnaya zadacha sprosa i predlozheniya [Transport demand and supply problem]. ITNOU, Informatsionnye tekhnologii v nauke, obrazovanii i upravlenii [Information technologies in science, education and management], 2018, no. 6 (10). pp. 83-88.
2. Zaychenko Yu.P. Issledovanie operatsiy: ucheb. dlya vuzov [Operations research], Kiev, Vyshcha shk. [Kyiv, Higher school], 1988, 552 p.
3. Hamdy A. Taha Vvedeniye v issledovaniye operatsiy [Introduction to operations research]. Moscow, Mir, 1988, 912 p.
4. Petrunin S.V. Resheniye transportnykh zadach PS-metodom pri ogranicheniyakh na peremennyye [Solution of transportation problems with bounded-variables using PC-method]. Nauchnyy vestnik MGTU GA [Scientific bulletin of the MSTU CA], 2014, no.202, pp. 53-57.
5. Shapkin A.S., Shapkin V.A. Matematicheskie metody i modeli issledovaniya operatsiy [Mathematical methods and models of operations research]. 7-ye izd. M., Izdatel'sko-torgovaya korporatsiya "Dashkov i Ko" [7 publ., Moscow, Publ. trading corporation "Dashkov i Ko"], 2019, 398 p.
6. Zolotaryuk A.V. Matematicheskaya model' mnogokriterial'noy optimizatsii transportnykh perevozok [Mathematical model of multi-criteria optimization of transport transportation]. Innovatsionnye tekhnologii v nauke i obrazovanii [Innovative technologies in science and education], 2015, no. 1(1). pp. 317-320
7. Nechitaylo N.M. Zadachi transportnogo tipa po kriteriyu vremeni s uchetom kharakteristik primenyaemykh transportnykh sredstv [Transport type tasks based on the time criterion, taking into account the characteristics of the transport vehicles used]. Mir transporta [World of transport and transportation], 2021, vol. 19, no. 3 (94), pp. 74-80.

8. Kubil V.N., Chernyshev Yu.O. Obzor dinamicheskikh zadach marshrutizatsii transporta [A review of dynamic vehicle routing problems]. Programmnyye produkty i sistemy [Software & Systems], 2020, vol. 33, no. 3, pp. 491–501.
9. Kostyuk Yu.L., Pozhidaev M.S. Priblizhennyye algoritmy resheniya sbalansirovannoy zadachi k kommivoyazherov [Approximate algorithms for solving the balanced k traveling salesman problem]. Vestnik Tomskogo gosudarstvennogo universiteta. Upravleniye, vychislitel'naya tekhnika i informatika [Bulletin of Tomsk state university. Management, computing and information science], 2008, no. 1(2), pp. 106–112.
10. Pozhidaev M.S., Kostyuk Yu.L. Sbalansirovannaya evristika dlya resheniya zadachi marshrutizatsii transporta s uchedom gruzopod"emnosti [Balanced heuristics for solving the problem of transport routing taking into account carrying capacity]. Vestnik TGU. UVTII [Bulletin TSU. MC&IS], 2010, no. 3, pp. 65-72.
11. Tyurin A.Yu. Utilization of advanced methods of food industry goods transportation routing. Vestn. KuzGTU [Bulletin of Kuzbass state technical university], 2011, no.4, pp.89-92.
12. Meindl B., Templ M. Analysis of commercial and free and open source solvers for linear optimization problems. Institut f. Statistik u. Wahrscheinlichkeitstheorie, Austria, 2012, 13 p.

Malanova Tatyana Valeryevna. Candidate of technical sciences, associate professor at software engineering center, institute of information technology and data analysis, Irkutsk national research technical university, AuthorID: 530825, ORCID: 0000-0002-1586-2530, malanova_tanya@mail.ru, 664074, Russian Federation, Irkutsk, Lermontov Str., 83.

Bakhvalov Sergey Vladimirovich. Candidate of technical sciences, associate professor at software engineering center, institute of information technology and data analysis, Irkutsk national research technical university, AuthorID: 169991, ORCID:0009-0009-83961926, bsv@istu.edu, 664074, Russian Federation, Irkutsk, Lermontov Str., 83.

Yanhaev Danil Vavimovich. Master's student, institute of information technology and data analysis, Irkutsk national research technical university, yanhaev2000@gmail.com, 664074, Russian Federation, Irkutsk, Lermontov Str., 83.

Arshinsky Vadim Leonidoich. Candidate of technical sciences, associate professor at software engineering center, institute of information technology and data analysis, Irkutsk national research technical university, AuthorID: 520305, ORCID: 0000-0001-7832-1582, arshinskyv@istu.edu, 664074, Russian Federation, Irkutsk, Lermontov Str., 83.

Статья поступила в редакцию 03.04.2024; одобрена после рецензирования 16.09.2024; принята к публикации 08.10.2024.

The article was submitted 04/03/2024; approved after reviewing 09/16/2024; accepted for publication 10/08/2024.

УДК 519.23:574.34:632.727:631.559

DOI:10.25729/ESI.2024.35.3.014

Параметрическая оптимизация производства аграрной продукции в условиях влияния на урожай саранчовых вредителей

Иваньо Ярослав Михайлович, Колокольцева Ирина Михайловна

Иркутский государственный аграрный университет имени А.А. Ежевского, Россия, Иркутск, iasa_econ@rambler.ru

Аннотация. Для решения задач оптимизации производственных объёмов растениеводческой продукции с учетом влияния на урожай саранчовых вредителей разработана модель параметрического программирования с вероятностными характеристиками. В качестве параметров определены температуры воздуха и осадки в начальный период вегетации, влияющие на численность саранчовых вредителей. Выделены три варианта математической модели: отсутствие влияния саранчовых на доход сельскохозяйственной организации; различное влияние саранчовых на экономический показатель; максимально возможное влияние вредителей с вероятностной оценкой ситуации. Предложены алгоритмы решения экстремальных задач. Модели и алгоритмы апробированы на данных сельскохозяйственной организации ЗАО «Иркутские семена» Иркутского района.

Ключевые слова: параметрическая оптимизация, факторная модель, саранчовые вредители, аграрная продукция, потери, метод статистических испытаний

Цитирование: Иваньо Я.М. Параметрическая оптимизация производства аграрной продукции в условиях влияния на урожай саранчовых вредителей / Я.М. Иваньо, И.М. Колокольцева // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С.157-166. – DOI: 10.25729/ESI.2024.35.3.014.

Введение. На урожай сельскохозяйственных культур влияет большое количество факторов: метеорологические (температура, осадки и др.), длительные бездождевые периоды, заморозки и т.д. Помимо этих явлений, риски значительных потерь могут быть связаны с биологическими событиями, одним из которых является нашествие вредителей. Поэтому теоретическую и практическую значимость имеют модели оптимизации с учётом потерь, вызванных вредителями. Особо следует выделить саранчовых вредителей, которые в наибольшей степени влияют на потери урожая, прежде всего, кормовых культур [1, 2].

Целью работы является разработка модели параметрического программирования для оптимизации объёмов производства растениеводческой продукции с учетом влияния на урожай саранчовых вредителей.

Для достижения цели решались следующие задачи:

- 1) разработка моделей для оценки влияния метеорологических факторов на численность саранчовых вредителей и урожайность сельскохозяйственных культур;
- 2) построение модели параметрического программирования для оптимизации объёмов производства растениеводческой продукции с учетом влияния на урожай саранчовых вредителей применительно к реальному объекту.

Методы и материалы. Для решения поставленных задач использованы методы теории вероятностей и математической статистики, параметрическое программирование.

В работе проведен корреляционно-регрессионный анализ [3] по данным о численности личинок саранчовых вредителей (экз./м²) и метеорологических факторах (температуры воздуха и осадков) за 2017-2023 гг. на примере муниципальных районов Иркутской области, а также зависимости урожайности сельскохозяйственных культур от численности личинок саранчовых вредителей.

Для оценки влияния саранчовых вредителей на урожайность разработан алгоритм с использованием метода Монте-Карло, который широко применяется для решения разных задач [4, 5], в том числе в сельском хозяйстве [6, 7, 8].

Сведения о численности вредителей сельскохозяйственных культур на территории Иркутской области получены по данным фитосанитарного мониторинга посевов сельскохозяйственных культур, проведённого специалистами ФГБУ «Российский сельскохозяйственный центр» за период 2017-2023 гг. [9-15].

В качестве исходной информации для оценки влияния вредителей на сельскохозяйственные растения на территории Иркутской области использованы многолетние ряды среднесуточных температур воздуха ($^{\circ}\text{C}$) и осадков (мм) за вегетационный период, полученные из фондов ФГБУ «Иркутское управление по гидрометеорологии и мониторингу окружающей среды».

Сведения по урожайности сельскохозяйственных культур в муниципальных районах Иркутской области получены по данным Территориального органа Федеральной службы государственной статистики по Иркутской области.

Для апробации разработанной модели использованы данные бухгалтерских отчётов ЗАО «Иркутские семена».

Основные результаты. Численность саранчовых зависит от таких факторов, как температура воздуха и осадки [16]. Пространственный анализ показывает, что метеорологические факторы на разных территориях Иркутской области влияют одинаково на результативный признак – прослеживается прямая зависимость от температуры и обратная – от осадков за июнь. Согласно регрессионному анализу, выявлены зависимости между численностью личинок саранчовых (экз./м²) y_i и факторами – средней месячной температурой воздуха T_i и суммой осадков x_i за этот же месяц. Для других месяцев вегетационного периода значимых регрессионных выражений не выявлено.

Для шести районов из восьми получены значимые регрессионные выражения, характеризующие связь результативного признака и метеорологических факторов за июнь [16]. Исключение составляют Иркутский и Черемховский районы. Здесь коэффициенты детерминации находятся ниже 0,50 (таблица 1).

В Тайшетском районе основным фактором, влияющим на численность саранчовых, являются суммы осадков за июль, хотя и для июня имеет место удовлетворительная связь. Между тем влияние саранчовых в этом районе на урожайность сельскохозяйственных культур ниже экономического порога вредоносности (ЭПВ) [17], поэтому полученное выражение не имеет смысла применять на практике.

Саранчовые вредители, как правило, повреждают все овощные, зерновые, пропашные и многие другие культуры [18, 19], но наибольшую опасность они причиняют сенокосам и пастбищам.

При наличии значимых регрессионных выражений, описывающих зависимости урожайности от численности вредителей, для решения задач по оптимизации объёмов производства растениеводческой продукции с учетом влияния на урожай саранчовых вредителей [1] можно использовать задачу параметрического программирования с вероятностными характеристиками. В этой задаче параметром является численность саранчовых, которая зависит от температуры воздуха и осадков в начальный период вегетации.

В этой задаче целевая функция описывает максимум дохода сельскохозяйственной организации:

$$\sum_{s \in S} c_s x_s \rightarrow \max, \quad (1)$$

Таблица 1. Регрессионные зависимости численности саранчовых от температур воздуха и сумм осадков за июнь в муниципальных районах Иркутской области по данным 2017-2023 гг. [16]

№ п.п.	Район	Выражение	R^2	F-критерий Фишера	Уровень значимости	t-статистика
1	Качугский	$y_t = 16,1T_t - 241,3$	0,80	20,00	0,0066	4,47
2	Тайшетский	$*y_t = 3,77 - 0,0269x_t$	0,75	12,5	0,0239	-3,54
3	Заларинский	$y_t = 12,08T_t - 183,3$	0,56	6,41	0,0523	2,53
4	Черемховский	$y_t = 5,12T_t - 0,061x_t - 61,7$	0,38	1,25	0,378	1,03; -0,39
5	Иркутский	$y_t = 5,29T_t - 80,5$	0,33	2,51	0,174	1,59
6	Боханский	$y_t = 6,17T_t - 96,1$	0,81	21,8	0,0055	4,67
7	Нукутский	$y_t = 22,8 - 0,142x_t$	0,65	9,31	0,0284	-3,05
8	Эхирит-Булагатский	$y_t = -120,8 + 7,62T_t$	0,67	8,29	0,0450	2,88

Примечание* Численность личинок саранчовых связана с осадками за июль при условиях:

- ограниченности производственных ресурсов

$$\sum_{s \in S} a_{ls} x_s \leq A_l, l \in L, \quad (2)$$

- ограниченности размера растениеводческой отрасли

$$\underline{n}_e \leq \sum_{s \in S} (1 + \beta_s) x_s \leq \bar{n}_e, e \in E, \quad (3)$$

- производства конечной продукции заданного объема

$$\sum_{s \in S} v_{qs}(y) x_s \geq V_q, q \in Q, \quad (4)$$

- определенного количества вносимых удобрений и средств защиты растений

$$\sum_{s \in S} b_{ms} x_s \leq B_m, m \in M, \quad (5)$$

- неотрицательности переменных

$$x_l \geq 0 \quad (6)$$

где x_s – искомая переменная, площадь сельскохозяйственной культуры s ; c_s – доход от производства единицы s -культуры; a_{ls} – расход ресурса l на единицу площади культуры s или вида кормовых угодий; L – множество ресурсов; A_l – наличие ресурса l -вида; V_q – гарантированный (обязательный объем) производства продукции вида q ; \bar{n}_l , \underline{n}_l – максимальная и минимальная возможная площадь культур группы e ; E – множество групп культур; $v_{qs}(y)$ – выход товарной продукции q -вида с единицы площади культуры s ; Q – множество видов продукции; β_s – коэффициент, учитывающий площадь семенных посевов для культуры s ; b_{ms} – расход удобрений видов m и средств защиты на единицу площади культуры s или вида кормовых угодий; B_m – необходимый объем удобрений вида m ; M – множество видов удобрений; y – параметр, характеризующий численность личинок саранчовых вредителей. При этом значение y изменяется в интервале $[\alpha, \gamma]$.

В этой модели:

$$v_{qs} = k_{0qs} + k_{1qs} y_{qs}, \quad (7)$$

где k_{0qs} , k_{1qs} – коэффициенты линейного уравнения регрессии.

С помощью модели (1) – (7) можно решить следующие задачи.

1. Влияние саранчовых вредителей на урожайность сельскохозяйственных культур практически отсутствует, товаропроизводители работают в усредненных условиях. Тогда неравенство (4) примет вид:

$$\sum_{s \in S} v_{qs} x_s \geq V_q, q \in Q. \quad (8)$$

2. Численность личинок саранчовых вредителей превышает экономический порог вредоносности (ЭПВ) [3]. В этом случае в модель (1) – (7) вводится дополнительное ограничение:

$$y_{qs} > \text{ЭПВ}. \quad (9)$$

3. Выделяется ситуация наибольшего влияния вредителей на урожайность сельскохозяйственных культур (наиболее благоприятные сочетания температур воздуха и осадков для роста численности личинок саранчовых вредителей). Для такой модели предлагается алгоритм вероятностной оценки редкого события.

Например, модель (1) – (9) реализована применительно к сельскохозяйственной организации ЗАО «Иркутские семена» Иркутского района (табл. 2), специализирующемся на производстве растениеводческой продукции.

В качестве переменных $x_1, x_2, x_3, x_4, x_5, x_6$ рассмотрены площади производства сельскохозяйственных культур: пшеница, ячмень, овес, картофель, многолетние травы и рапс.

Исследования показали, что наибольшее влияние саранчовые вредители оказывают на урожай пшеницы и многолетних трав, поэтому в неравенстве (4) рассмотрены вероятные потери урожайности этих культур.

Моделировать численность саранчовых вредителей можно двумя способами при наличии границ интервала. В первом из них последовательно увеличивается значение y_{qs} от нижней до верхней оценки. Во втором случае используется метод статистических испытаний, случайным образом моделируется численность саранчовых вредителей в заданных интервалах. Таким образом, оптимальные решения определяются многократно в зависимости от численности саранчовых вредителей.

Применение метода статистических испытаний адекватно может отражать конкретную ситуацию, поскольку численность саранчовых вредителей изменяется непредсказуемо, как сочетание осадков и температур воздуха в начале вегетации. Первый же способ позволяет оценить разные ситуации, связанные с потерями доходов.

Выделены оптимальные планы, соответствующие усреднённым условиям работы товаропроизводителей, наименьшему, наибольшему и медианному значению доходов, при условии превышения экономического порога вредоносности. Для получения результатов реализован алгоритм с использованием метода Монте-Карло.

Анализ результатов решения задачи показал снижение доходности по растениеводческой отрасли с учётом влияния саранчовых вредителей на 4,3% по сравнению с усреднёнными условиями, что соответствует 10,6 млн рублей. Потери объёмов производства составят 6,6% или 981 т, из которых 882 т приходится на пшеницу, 99 т – на многолетние травы.

В табл. 2 также приведены результаты моделирования редких событий, когда саранчовые вредители в значительной степени влияют на потери урожая. Для оценки редкого события (маловероятной численности саранчовых вредителей) использовалась линейная факторная зависимость:

$$y_t = -100,9 + 6,3T_t, \quad (10)$$

где T_t – средняя многолетняя суточная температура воздуха с 7 по 26 июня. Выражение (10) значимо для уровня значимости F -критерия Фишера 0,024.

Таблица 2. Результаты решения задачи параметрического программирования для ЗАО «Иркутские семена»

Показатель	Численность личинок, экз/м ²	Оптимальные планы						Целевая функция, млн руб.
		Пшеница	Ячмень	Овес	Картофель	Многолетние травы	Рапс	
Усреднённые условия								
Площади, га	0,1	x_1	x_2	x_3	x_4	x_5	x_6	248,3
Урожайность, т/га		2100	486	100	560	76	10	
Объемы производства, т		2,23	1,90	1,70	16,00	2,00	0,70	
Неблагоприятные условия:								
Нижняя оценка								
Площади, га	81,2	x_1	x_2	x_3	x_4	x_5	x_6	237,7
Урожайность, т/га		2100	486	100	560	76	10	
Объемы производства, т		1,81	1,90	1,70	16,00	0,64	0,70	
Медиана								
Площади, га	37,8	x_1	x_2	x_3	x_4	x_5	x_6	243,5
Урожайность, т/га		2100	486	100	560	76	10	
Объемы производства, т		2,04	1,90	1,70	16,00	0,75	0,70	
Верхняя оценка								
Площади, га	10,1	x_1	x_2	x_3	x_4	x_5	x_6	247,0
Урожайность, т/га		2100	486	100	560	76	10	
Объемы производства, т		2,18	1,90	1,70	16,00	1,63	0,70	
Редкое событие								
Площади, га	120,0	x_1	x_2	x_3	x_4	x_5	x_6	232,6
Урожайность, т/га		2100	486	100	560	76	10	
Объемы производства, т		1,61	1,90	1,70	16,00	1,89	0,70	
Объемы производства, т		3381	923	170	8960	144	7	

Моделирование ряда среднемесячных температур воздуха T_i осуществлялось с помощью закона распределения вероятностей Пирсона III типа [20, 21] со статистическими оценками, полученными по данным 1971-2023 гг. (среднее – 16,2°C, коэффициенты вариации и асимметрии – 0,11 и -0,18) (иллюстрируется на рисунке 1). По эмпирическим данным, наибольшее значение T_i наблюдалось в 2000 году и составило 19,9°C. Для этого значения температуры воздуха численность саранчовых с вероятностью 0,0346 соответствует 24,5 экз./м².

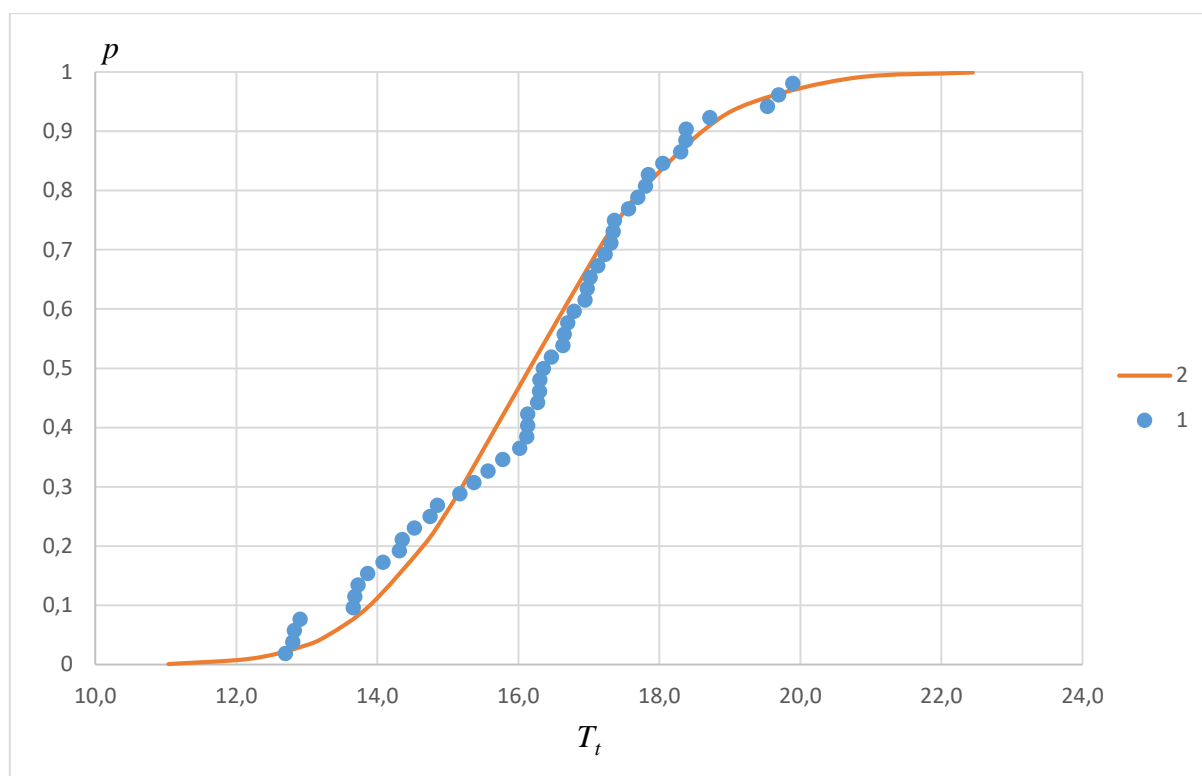


Рис. 1. Эмпирическая функция (1) и распределение Пирсона III типа (2) для средних многолетних температур воздуха с 7 по 26 июня по данным г. Иркутска за 1971-2023 гг.

Однако на некоторых участках заброшенных земель, пастбищах и отдельных хозяйствах, где не проводятся защитные и агротехнические мероприятия, численность вредителя значительно выше. По области за многолетний период зафиксирована численность саранчовых, превышающая 300 экз./м², а в Иркутском районе – 120 экз./м². Поэтому в данной работе смоделирована ситуация проявления редкого события – очень высокой численности вредителя.

Для оценки вероятности такого события с помощью псевдослучайных чисел многократно моделируется ряд численности вредителей в заданном интервале (10,0 – 120,0 экз./м²) до появления значения, которое соответствует фактическому редкому явлению (120 экз./м²). Определяется расчётная эмпирическая вероятность редкого события по формуле $p = t / (n + 1)$, где t – порядковый номер, n – объём выборки [20]. Эксперименты повторяются с моделированием k выборок. Моделирование k выборок позволяет получить такое же количество вероятностей появления редкого события. Чем большее число выборок моделируется, тем точнее статистические оценки вероятностей при условии стационарности изменчивости событий. Среднее значение k вероятностей p_c характеризует формирование редкого события. Для конкретного случая вероятность p_c при $k = 100$ равна 0,0153.

Для редкого события применительно к ЗАО «Иркутские семена» рассчитано возможное снижение дохода на 6,3 % по сравнению с усреднёнными условиями, что соответствует 15,7 млн рублей.

Заключение. Предложена модель параметрического программирования с параметром в виде численности саранчовых вредителей, который зависит от метеорологических факторов – средней температуры воздуха и осадков в начальный период вегетации. Определены значимые линейные зависимости урожайности зерновых и многолетних трав от численности личинок саранчовых вредителей.

Выделены три варианта математической модели: отсутствие влияния вредителей (задача линейного программирования); численность саранчовых вредителей выше экономического

порога вредоносности (параметрическая задача с интервальными оценками); наиболее неблагоприятные условия получения продукции (модель с вероятностной оценкой редкого события). Приведен алгоритм оценки маловероятной численности саранчовых вредителей и условий её проявления.

Модели и алгоритмы реализованы для ЗАО «Иркутские семена» Иркутского района с оценкой вероятных потерь дохода и продукции от саранчовых вредителей.

Благодарности. Работа выполнена при поддержке Российского научного фонда, проект № 24-21-00502.

Список источников

1. Барсукова М.Н. О некоторых моделях оптимизации производства растениеводческой продукции в условиях биологических рисков / М.Н. Барсукова, Я.М. Иванько, И.М. Колокольцева // Моделирование систем и процессов, 2022. – Т. 15. – № 4. – С. 17-23. – DOI:10.12737/2219-0767-2022-15-4-17-23. – EDN NBGJHS.
2. Иванько Я.М. Моделирование аграрного производства в условиях влияния климатических, экологических и биологических событий / Я.М. Иванько, И.М. Колокольцева, С.А. Петрова // System analysis and mathematical modeling, 2023. – Т. 5. – № 4. – С. 442-456. – DOI:10.17150/2713-1734.2023.5(4).442-456. – EDN XPFQBV.
3. Ферстер Э. Методы корреляционного и регрессионного анализа: Руководство для экономистов / Э. Ферстер, Б. Ренц; Пер. с нем. и предисл. В.М. Ивановой. – Москва: Финансы и статистика, 1983. – 302 с.
4. Robert C.P., George C. Monte Carlo statistical methods. Springer texts in statistics, 2005, DOI:10.1007/978-1-4757-4145-2.
5. Фалин Г.И. Анализ рисков с помощью метода Монте-Карло. / Г.И. Фалин // Управление риском, 2017. – № 1. – 20 с.
6. Buzina T.S., Belyakova A.Yu., Ivanyo Ya.M. Method of statistical tests in solving problems of food production management. IOP Conference series: Earth and environmental science. Krasnoyarsk, IOP Publishing Ltd, 2021, vol. 839, p. 32051, DOI:10.1088/1755-1315/839/3/032051, EDN KYPGSP.
7. Иванько Я.М. Метод статистических испытаний в моделировании урожайности зерновых культур в зависимости от метеорологических факторов / Я.М. Иванько, С.А. Петрова // Актуальные вопросы инженерно-технического и технологического обеспечения АПК : Материалы VIII Национальной научно-практической конференции с международным участием «Чтения И.П. Терских», посвященной 85-летию Иркутского ГАУ, Иркутск, 26-27 сентября 2019 года. – Иркутск: Иркутский государственный аграрный университет имени А.А. Ежевского, 2019. – С. 174-179. – EDN PBMТMT.
8. Решение задач управления аграрным производством в условиях неполной информации / Я.М. Иванько, П.Г. Асалханов, М.Н. Полковская и др. // Министерство сельского хозяйства Российской Федерации ФГБОУ ВПО Иркутская государственная сельскохозяйственная академия. – Иркутск: Иркутская государственная сельскохозяйственная академия, 2012. – 200 с. – ISBN 978-5-91777-085-7. – EDN BPAABQ.
9. Обзор фитосанитарного состояния посевов сельскохозяйственных культур в Российской Федерации в 2017 году и прогноз развития вредных объектов в 2018 году / Филиал ФГБУ Российский сельскохозяйственный центр по Иркутской области, 2018.– М..
10. Обзор фитосанитарного состояния посевов сельскохозяйственных культур в Российской Федерации в 2018 году и прогноз развития вредных объектов в 2019 году / Филиал ФГБУ Российский сельскохозяйственный центр по Иркутской области, 2019 – М..
11. Обзор фитосанитарного состояния посевов сельскохозяйственных культур в Российской Федерации в 2019 году и прогноз развития вредных объектов в 2020 году / Филиал ФГБУ Российский сельскохозяйственный центр по Иркутской области, 2020. – М..
12. Обзор фитосанитарного состояния посевов сельскохозяйственных культур в Российской Федерации в 2020 году и прогноз развития вредных объектов в 2021 году / Филиал ФГБУ Российский сельскохозяйственный центр по Иркутской области, 2021. – М..
13. Обзор фитосанитарного состояния посевов сельскохозяйственных культур в Российской Федерации в 2021 году и прогноз развития вредных объектов в 2022 году / Филиал ФГБУ Российский сельскохозяйственный центр по Иркутской области, 2022.– М..
14. Обзор фитосанитарного состояния посевов сельскохозяйственных культур в Российской Федерации в 2022 году и прогноз развития вредных объектов в 2023 году / Филиал ФГБУ Российский сельскохозяйственный центр по Иркутской области, 2023. – М., 2023.

15. Обзор фитосанитарного состояния посевов сельскохозяйственных культур в Российской Федерации в 2023 году и прогноз развития вредных объектов в 2024 году / Филиал ФГБУ Российский сельскохозяйственный центр по Иркутской области, 2024. – М..
16. Колокольцева И.М. Оценка влияния метеорологических факторов на численность саранчовых / И.М. Колокольцева, Я.М. Иваньо // Материалы международной научно-практической конференции молодых ученых «Научные исследования и разработки к внедрению в АПК». – Молодежный: Изд-во Иркутский ГАУ, 2024. – С.105-110.
17. Алехин В.Т. Экономические пороги вредоносности вредителей, болезней и сорных растений в посевах сельскохозяйственных культур: справочник / В.Т. Алехин, В.В. Михайликова, Н.Г. Михина. – Москва: Росинформагротех, 2016.
18. Колокольцева И.М. Группирование вредителей и болезней сельскохозяйственных культур по влиянию на ущерб / И.М. Колокольцева, Я.М. Иваньо // Проблемы и перспективы устойчивого развития агропромышленного комплекса : Материалы Всероссийской научно-практической конференции с международным участием, посвященной памяти Александра Александровича Ежевского, п. Молодежный, 16-17 ноября 2023 года. – п. Молодежный: Иркутский государственный аграрный университет имени А.А. Ежевского, 2023. – С. 553-559. – EDN WWOQLH.
19. Митюшев И.М. Вредители сельскохозяйственных растений / И. М. Митюшев // Научно-образовательный портал "Большая российская энциклопедия", 2023. – № 4. – DOI 10.54972/00000001_2023_4_33. – EDN DDMIBJ.
20. Крицкий С.Н. Гидрологические основы управления водохозяйственными системами / С.Н. Крицкий, М.Ф. Менкель. – М.: Наука, 1982. – 271 с.
21. Справочник по теории вероятностей и математической статистике / В.С. Королюк, Н.И. Портенко, А.В. Скороход, А.Ф. Турбин. – 2-е изд., перераб. и доп. – Москва: Наука, 1985. – 640 с.; 21 см.

Иваньо Ярослав Михайлович. Доктор технических наук, профессор, проректор по цифровой трансформации, ФГБОУ ВО Иркутский ГАУ, AuthorID: 106394, SPIN: 9654-8057, ORCID ID: 0000-0003-4118-7185, iasa_econ@rambler.ru, 664038, Россия, Иркутская область, Иркутский район, п. Молодёжный.

Колокольцева Ирина Михайловна. Аспирант кафедры информатики и математического моделирования, ФГБОУ ВО Иркутский ГАУ, AuthorID: 1109185, SPIN: 9426-2093, ORCID ID: 0009-0008-5653-1800, 89025190281@yandex.ru, 664038, Россия, Иркутская область, Иркутский район, п. Молодёжный.

UDC 519.23:574.34:632.727:631.559

DOI:10.25729/ESI.2024.35.3.014

Parametric optimization of agricultural production under the influence of locust pests on the harvest

Yaroslav M. Ivanyo, Irina M. Kolokoltseva

Irkutsk state agricultural university named after A.A. Yezhevsky,
Russia, Irkutsk, *iasa_econ@rambler.ru*

Abstract. To solve the problems of optimizing the production volumes of crop production, taking into account the impact of locust pests on the harvest, a parametric programming model with probabilistic characteristics has been developed. Air temperatures and precipitation during the initial growing season, which affect the number of locust pests, were determined as parameters. Three variants of the mathematical model are distinguished: the absence of the influence of locusts on the income of an agricultural organization; the different influence of locusts on the economic indicator; the maximum possible impact of pests with a probabilistic assessment of the situation. Algorithms for solving extreme problems are proposed. The models and algorithms have been tested on the data of the agricultural organization CJSC Irkutsk Seeds of the Irkutsk region.

Keywords: parametric optimization, factor model, locust pests, agricultural products, losses, statistical testing method

References

1. Barsukova M.N., Ivanyo Ya.M., Kolokoltseva I.M. O nekotoryh modelyakh optimizacii proizvodstva rastenievodcheskoj produkcii v usloviyah biologicheskikh riskov [On some models of optimization of crop production in conditions of biological risks]. Modelirovanie sistem i processov [Modeling of systems and processes], 2022, vol. 15, no. 4, pp. 17-23, DOI 10.12737/2219-0767-2022-15-4-17-23.
2. Ivanyo Ya.M., Kolokoltseva I.M., Petrova S.A. Modelirovaniye agrarnogo proizvodstva v usloviyakh vliyaniya klimaticheskikh, ekologicheskikh i biologicheskikh sobytii [Modeling of agricultural production under the influence of climatic, ecological and biological events]. System analysis and mathematical modeling [System analysis and mathematical modeling], 2023, vol. 5, no. 4, pp. 442-456, DOI:10.17150/2713-1734.2023.5(4).
3. Ferster E., Renz B. Metody korrelyacionnogo i regressionnogo analiza [Methods of correlation and regression analysis]. Finansy i statistika [Moscow, Finance and Statistics], 1983, 304 p.
4. Robert C.P., George C. Monte Carlo statistical methods. Springer texts in statistics, 2005, DOI:10.1007/978-1-4757-4145-2.
5. Falin G.I. Analiz riskov s pomoshch'yu metoda Monte-Karlo [Risk analysis using the Monte Carlo method]. Upravlenie riskom [Risk management], 2017, no. 1, 20 p.
6. Buzina T.S., Belyakova A.Yu., Ivanyo Ya.M. Method of statistical tests in solving problems of food production management. IOP Conference series: Earth and environmental science. Krasnoyarsk, IOP Publishing Ltd, 2021, vol. 839, p. 32051, DOI:10.1088/1755-1315/839/3/032051, EDN KYPGSP.
7. Ivanyo Ya.M., Petrova S.A. Metod statisticheskikh ispytaniy v modelirovanii urozhajnosti zernovykh kul'tur v zavisimosti ot meteorologicheskikh faktorov [The method of statistical tests in modeling the yield of grain crops depending on meteorological factors]. Aktual'nye voprosy inzhenerno-tekhnicheskogo i tekhnologicheskogo obespecheniya APK. Molodyozhnyj: Irkutskij gosudarstvennyj agrarnyj universitet imeni A.A. Ezhevskogo [Actual issues of engineering, technical and technological support of the agro-industrial complex: Proceedings of the VIII National scientific and practical conference with international participation "Readings of I.P. Terskikh", dedicated to the 85th anniversary of Irkutsk State Agrarian University, Irkutsk, September 26-27, 2019, Irkutsk: Irkutsk State Agrarian University named after A.A. Ezhevsky], 2019, pp. 174-179.
8. Ivanyo Ya.M., Asalkhanov P.G., Polkovskaya M.N. et al. Reshenie zadach upravleniya agrarnym proizvodstvom v usloviyah nepolnoj informacii [Solving problems of agricultural production management in conditions of incomplete information]. Irkutsk, Irkutskaya gosudarstvennaya sel'skohozyajstvennaya akademiya [Irkutsk, Irkutsk state agricultural academy], 2012, 200 p.
9. Obzor fitosanitarnogo sostoyaniya posevov sel'skohozyajstvennykh kul'tur v Rossijskoj Federacii v 2017 godu i prognoz razvitiya vrednykh ob"ektov v 2018 godu. [Overview of the phytosanitary condition of crops in the Russian Federation in 2017 and forecast of the development of harmful objects in 2018]. Filial FGBU Rossijskij sel'skohozyajstvennyj centr po Irkutskoj oblasti, 2018.
10. Obzor fitosanitarnogo sostoyaniya posevov sel'skohozyajstvennykh kul'tur v Rossijskoj Federacii v 2018 godu i prognoz razvitiya vrednykh ob"ektov v 2019 godu. [Overview of the phytosanitary condition of crops in the Russian Federation in 2018 and forecast of the development of harmful objects in 2019]. Filial FGBU Rossijskij sel'skohozyajstvennyj centr po Irkutskoj oblasti, 2019.
11. Obzor fitosanitarnogo sostoyaniya posevov sel'skohozyajstvennykh kul'tur v Rossijskoj Federacii v 2019 godu i prognoz razvitiya vrednykh ob"ektov v 2020 godu. [Overview of the phytosanitary condition of crops in the Russian Federation in 2019 and forecast of the development of harmful objects in 2020]. Filial FGBU Rossijskij sel'skohozyajstvennyj centr po Irkutskoj oblasti, 2020.
12. Obzor fitosanitarnogo sostoyaniya posevov sel'skohozyajstvennykh kul'tur v Rossijskoj Federacii v 2020 godu i prognoz razvitiya vrednykh ob"ektov v 2021 godu. [Overview of the phytosanitary condition of crops in the Russian Federation in 2020 and forecast of the development of harmful objects in 2021]. Filial FGBU Rossijskij sel'skohozyajstvennyj centr po Irkutskoj oblasti, 2021.
13. Obzor fitosanitarnogo sostoyaniya posevov sel'skohozyajstvennykh kul'tur v Rossijskoj Federacii v 2021 godu i prognoz razvitiya vrednykh ob"ektov v 2022 godu. [Overview of the phytosanitary condition of crops in the Russian Federation in 2021 and forecast of the development of harmful objects in 2022]. Filial FGBU Rossijskij sel'skohozyajstvennyj centr po Irkutskoj oblasti, 2022.
14. Obzor fitosanitarnogo sostoyaniya posevov sel'skohozyajstvennykh kul'tur v Rossijskoj Federacii v 2022 godu i prognoz razvitiya vrednykh ob"ektov v 2023 godu. [Overview of the phytosanitary condition of crops in the Russian Federation in 2022 and forecast of the development of harmful objects in 2023]. Filial FGBU Rossijskij sel'skohozyajstvennyj centr po Irkutskoj oblasti, 2023.
15. Obzor fitosanitarnogo sostoyaniya posevov sel'skohozyajstvennykh kul'tur v Rossijskoj Federacii v 2023 godu i prognoz razvitiya vrednykh ob"ektov v 2024 godu. [Overview of the phytosanitary condition of crops in the Russian

- Federation in 2023 and forecast of the development of harmful objects in 2024]. Filial FGBU Rossijskij sel'skokhozyajstvennyj centr po Irkutskoj oblasti, 2024.
16. Kolokoltseva I.M., Ivanyo Ya.M. Ocenka vliyaniya meteorologicheskikh faktorov na chislennost' saranchovyh [Assessment of the influence of meteorological factors on the number of locusts]. Materialy mezhdunarodnoj nauchno-prakticheskoy konferencii molodyh uchenyh "Nauchnye issledovaniya i razrabotki k vnedreniyu v APK". Molodezhnyj, Irkutskij GAU [Proceedings of the international scientific and practical conference of young scientists "Scientific research and development for implementation in the agro-industrial complex" – Youth, Publishing house of Irkutsk State Agrarian University], 2024, pp. 105-110.
 17. Alyokhin V.T., Mikhaylikova V.V., Mikhina N.G. Ekonomicheskiye porogi vredonosnosti vreditel'nykh i bolezney i sornykh rasteniy v posevakh sel'skokhozyaystvennykh kultur: spravochnik [Economic thresholds of harmfulness of pests, diseases and weeds in crops: handbook]. Rosinformagrotekh, 2016.
 18. Kolokoltseva I.M., Ivanyo Ya.M. Gruppyrovaniye vreditel'nykh i bolezney sel'skokhozyaystvennykh kultur po vliyaniyu na ushcherby [Grouping of pests and diseases of agricultural crops by their effect on damage]. Problemy i perspektivy ustoychivogo razvitiya agropromyshlennogo kompleksa [Problems and prospects of sustainable development of the agro-industrial complex: Proceedings of the All-Russian scientific and practical conference with international participation dedicated to the memory of Alexander Alexandrovich Ezhevsky, p. Molodezhny, November 16-17, 2023, p. Molodezhny, Irkutsk state agrarian university named after A.A. Ezhevsky], 2023, pp. 553-559, EDN WWOQLH.
 19. Mityushev I.M. Vrediteli sel'skokhozyaystvennykh rasteniy [Pests of agricultural plants]. Nauchno-obrazovatelnyy portal "Bolshaya rossiyskaya entsiklopediya" [Scientific and educational portal "Great Russian Encyclopedia"], 2023, no. 4, DOI:10.54972/00000001_2023_4_33, EDN DDMIBJ.
 20. Kritsky S.N., Menkel M.F. Gidrologicheskiye osnovy upravleniya vodokhozyaystvennyimi sistemami [Hydrological foundations of water management systems]. Moscow, Nauka [Science], 1982, 271 p.
 21. Korolyuk V.S., Portenko N.I., Skorokhod A.V., Turbin A.F. [Handbook of probability theory and mathematical statistics] Moscow, Nauka [Science], 1985, 640 p.

Ivanyo Yaroslav Mikhailovich. Doctor of technical sciences, professor, vice-rector for digital transformation, Irkutsk state agricultural university, AuthorID: 106394, SPIN: 9654-8057, ORCID: 0000-0003-4118-7185, iasa_econ@rambler.ru, 664038, Russia, Irkutsk region, Irkutsk district, Molodezhny settlement.

Kolokoltseva Irina Mikhailovna. Postgraduate student of the department of computer science and mathematical modeling, Irkutsk state agricultural university, AuthorID: 1109185, SPIN: 9426-2093, ORCID: 0009-0008-5653-1800, 89025190281@yandex.ru, 664038, Russia, Irkutsk region, Irkutsk district, Molodezhny settlement.

Статья поступила в редакцию 02.09.2024; одобрена после рецензирования 10.09.2024; принята к публикации 10.10.2024.

The article was submitted 09/02/2024; approved after reviewing 09/10/2024; accepted for publication 10/10/2024.

УДК 537.52:519.624

DOI:10.25729/ESI.2024.35.3.015

Характеристики струйного высокочастотного индукционного разряда пониженного давления в процессах обработки твердых тел

Шемахин Александр Юрьевич¹, Абдуллин Ильдар Шаукатович², Желтухин Виктор Семенович¹, Макеева Елена Александровна¹

¹Казанский (Приволжский) федеральный университет,
Россия, Казань, ashemakhin@yandex.ru,

²НПО «Плазма ВСТ»

Аннотация. Проведены экспериментальные измерения параметров струйного высокочастотного индукционного разряда пониженного давления в разрядной камере и пламенной струе, измерения энергии ионов и плотности ионного тока на поверхности образца, а также проведена валидация математической модели струйного ВЧИ-разряда пониженного давления на экспериментальных данных. Установлено, что струйный ВЧИ-разряд пониженного давления в диапазоне давлений 13,3-133 Па является отдельной разновидностью высокочастотного разряда пониженного давления комбинированного типа, который отличается как от ВЧИ-разряда в потоке газа атмосферного давления с соленоидальным индуктором, так и от ВЧИ-разряда низкого давления в плазмотроне с плоской спиральной антенной, и от ВЧИ-разряда низкого давления с соленоидальным индуктором.

Ключевые слова: струйный ВЧИ-разряд, пониженное давление, поток плазмы, экспериментальные исследования, валидация модели

Цитирование: Шемахин А.Ю. Характеристики струйного высокочастотного индукционного разряда пониженного давления в процессах обработки твердых тел / А.Ю. Шемахин, И.Ш. Абдуллин, В.С. Желтухин, Е.А. Макеева // Информационные и математические технологии в науке и управлении, 2024. – № 3(35). – С. 167-175. – DOI: 10.25729/ESI.2024.35.3.015.

Введение. Высокочастотная индукционная (ВЧИ) плазма пониженного давления (13,3 – 133 Па) с продувом газа применяется для модификации различных материалов: диэлектриков, проводящих, полупроводниковых с целью изменения их качественных свойств, повышения прочности, надежности и долговечности изделий [1-2]. С каждым годом расширяется область применения ВЧ-плазмы [3-10]. Образованная ВЧ-индукционным разрядом плазма характеризуется следующими параметрами: концентрация электронов $n_e = 10^{15}-10^{18} \text{ м}^{-3}$, степень ионизации $10^{-6}-10^{-3}$, электронная температура $T_e = 1-4 \text{ эВ}$, температура атомов и ионов в плазменном сгустке $T_a = (3-4) \cdot 10^3 \text{ К}$, в струе $T_a = (3.2-10) \cdot 10^2 \text{ К}$.

Основными параметрами, определяющими результаты модификации твердых тел в «холодной» плазме струйного ВЧИ-разряда пониженного давления, являются энергия ионов и плотность ионного тока на поверхности образца. В данной работе проведены экспериментальные измерения основных параметров струйного ВЧИ-разряда пониженного давления и характеристик обработки – энергии ионов и ионного тока, поступающих на поверхность образца. Исследованы характеристики струйного ВЧИ-разряда индукционного типа с соленоидальным индуктором в атмосфере аргона в следующем диапазоне рабочих параметров: давление $p = 13,3 - 133 \text{ Па}$, расход газа $G \leq 0,4 \text{ г/с}$, потребляемая мощность генераторов $P_{\text{потр}}$ от 1,5 до 20 кВт, частота $f = 1,76 \text{ МГц}$.

В работах [11-15] описана конфигурация установки и построена математическая модель, разработаны численный метод и программный комплекс для расчета параметров струйного ВЧИ-разряда пониженного давления. В настоящей работе, в дополнение к экспериментальным исследованиям, проводится валидация программного комплекса по ряду экспериментальных данных. Далее будут рассмотрены зависимости полученных экспериментальных данных от мощности разряда и расхода газа, по пространственным распределениям в вакуумной и разрядной камерах и по параметрам обработки материалов различной физической природы.

1. Зависимости от мощности разряда. На рис. 1а приведены экспериментальные данные по мощности, выделяемой в ВЧ-разряде пониженного давления, измеренной колориметрическим методом в зависимости от мощности на индукторе. Разница составила не более 10% с расчетными данными.

Измерения распределений скорости в вакуумной камере показали, что скорость истечения плазмы на срезе плазматрона составляет от 250 до 450 м/с, что меньше скорости звука. Поэтому в области обрабатываемого тела не возникают ударные волны и не происходит резкого поднятия температуры потока, следовательно, роль плазмы сводится к транспортировке заряженных частиц и атомов к поверхности твердого тела и интенсификации тепло и массообмена (рис. 1б).

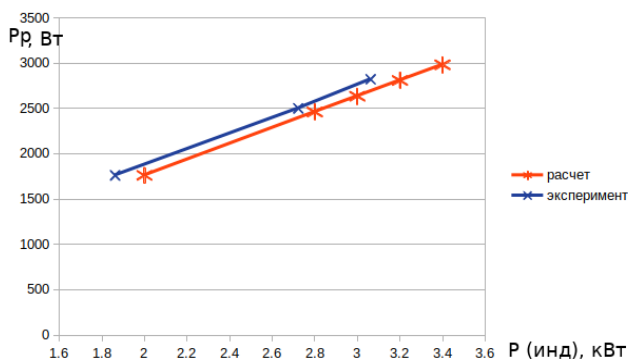


Рис. 1а. Валидация математической модели по экспериментальным данным мощности, выделяемой в разряде от мощности на индукторе

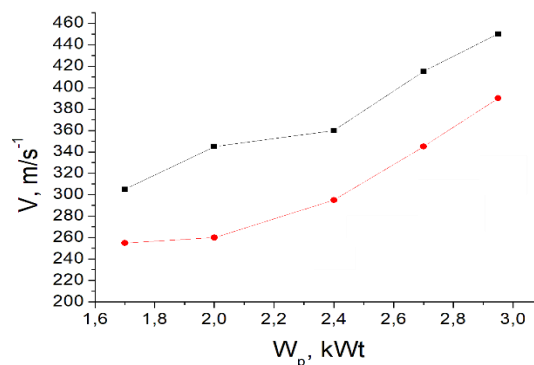


Рис. 1б. Зависимость скорости потока плазмы аргона на срезе разрядной камеры от мощности разряда, черная линия – при расходе 0,14 г/с, красная – при расходе 0,1 г/с

2. Зависимости от расхода плазмообразующего газа. Измерения распределений скорости в рабочей камере проведены с помощью трубки Пито. Из рис. 2а видно, что с увеличением расхода газа скорость плазменного потока растет приблизительно линейно.

В плазменной струе аксиальная компонента H_z относительно расхода ведет себя практически линейно (рис. 2б). Следовательно, наблюдается постоянство напряженности поля при разных расходах.

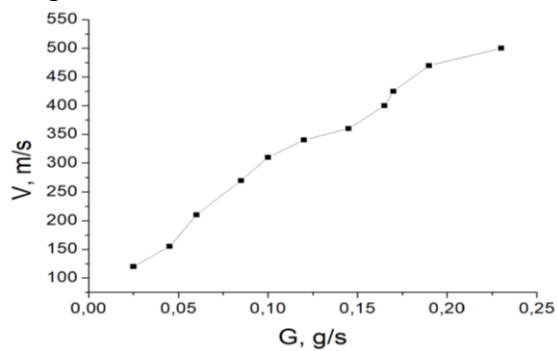


Рис. 2а. Зависимость скорости потока плазмы от расхода плазмообразующего газа при мощности разряда 2,4 кВт на срезе плазматрона на оси камеры.

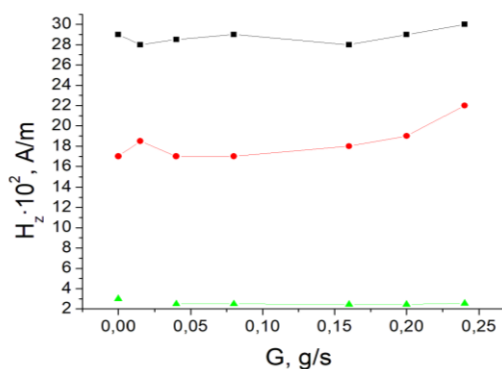


Рис. 2б. Зависимость напряженности магнитного поля в плазме ВЧИ-разряда от расхода плазмообразующего газа $P_p = 2,4$ кВт, $p = 113$ Па, $f = 1,76$ МГц: $z = -60$ мм черная линия – аргон $r = R$, красная линия – аргон $r = 0$; зеленая линия – $z = 120$ мм аргон $r = 0$.

3. Пространственное распределение характеристик струйного ВЧИ-разряда пониженного давления. Обнаружен эффект образования зоны перегрева по периферии струи

около входного отверстия вакуумной камеры, в котором температура потока на периферии струи выше температуры в центре струи на 50°C при скорости потока на входе в вакуумную камеру более 440 м/с и расходе плазмообразующего газа более $0,16\text{ г/с}$ при $2,4\text{ кВт}$, который подтвержден экспериментально с помощью термопар, расположенных на разном расстоянии от центра потока. Из рис. 3 видно, что на периферии струи наблюдается температурный пик, температура в центре потока выше на 51 градус , в то время как в расчетах получено превышение на $131\text{ градус Цельсия}$. В центре расчетная – 1231 К , на периферии – 1362 К , экспериментальная в центре – 1197 К , на периферии – 1248 градусов , то есть максимальное значение температуры создается на периферии струи, а локальный минимум температуры достигается в центре. Температура в поперечном сечении быстро спадает при удалении от струи. Точками обозначены экспериментальные данные (рис. 3), отклонение составляет не более 20% .

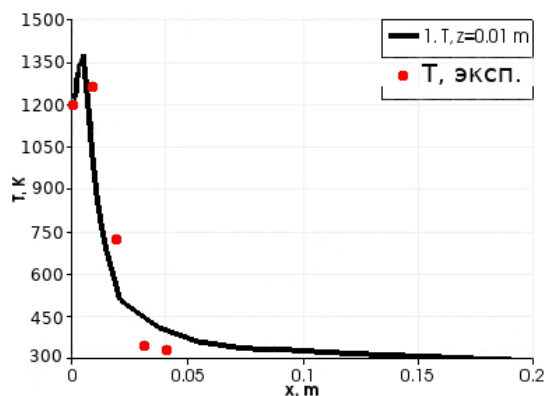


Рис. 3. Радиальное распределение температуры T_a (на рис. обозн. T [K]) в поперечном сечении струи ($Rrk = 0,012\text{ м}$, $Rvk = 0,2\text{ м}$, $Lvk = 0,5\text{ м}$, $G = 0,16\text{ г/с}$, $P=2,4\text{ кВт}$, $v = 440\text{ м/с}$). Точками обозначен эксперимент, подтверждающий наличие эффекта перегрева струи по периферии

Интерпретация этого эффекта следующая: в центре струи несущий газ охлаждается из-за резкого расширения потока при истечении из разрядной камеры; на периферии струи поток плазмы встречается с неподвижным газом внутри вакуумной камеры, при этом возрастает локальное давление газа, что приводит к увеличению частоты столкновений электронов с атомами и ионами, частоты столкновений возбужденных частиц; следствием этого является увеличение объемной рекомбинации заряженных частиц и интенсифицируются процессы гашения частиц в возбужденных состояниях. В процессах рекомбинации и гашения возбужденных состояний выделяется энергия, затраченная на ионизацию и возбуждение атомов. Совокупное воздействие газодинамических и плазмодинамических эффектов приводит к повышению температуры газа на периферии струи. Сопоставление расчетов с экспериментом является дополнительной валидацией разработанной модели.

С целью оценки влияния ВЧ-электромагнитного поля на обработку различных материалов, в том числе металлических, измерялась напряженность магнитного поля. Результаты измерений представлены на рис. 4.

В плазменной струе аксиальная компонента магнитного поля H_z не превышает 2350 А/м , а в сгустке плазмы в разрядной камере она достигает значений 5750 А/м (рис. 4а). В области сгустка находится максимум вдоль оси. Распределение магнитного поля по радиусу разрядной камеры при мощности разряда $W_p=2,4\text{ кВт}$ на расстоянии 120 мм от среза плазмотрона в сторону индуктора представлено на рис. 4б.

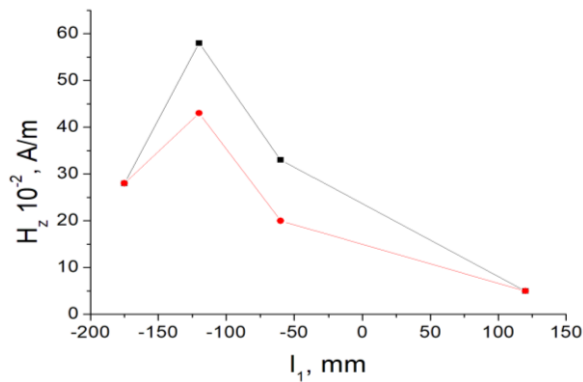


Рис. 4а. Распределение напряженности магнитного поля по оси плазматрона при мощности разряда $W_p=2,4$ кВт, черная линия — $r=R$, $G=0$ аргон; красная — $r=0$, $G=0,4$ г/с, аргон.

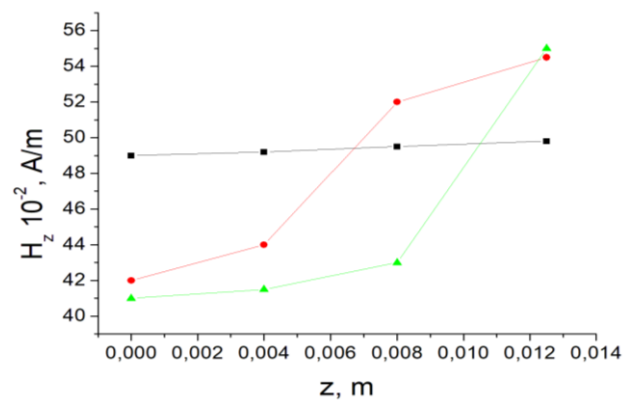


Рис. 4б. Распределение напряженности магнитного поля по радиусу разрядной камеры при мощности разряда $W_p=2,4$ кВт, $l_1=-120$ мм, черная линия — без плазмы; красная — аргон $p=113$ Па, $G=0,04$ г/с; зеленая — $p=113$ Па, $G=0$ г/с.

4. Взаимодействие струйного ВЧИ-разряда пониженного давления с материалами.

Свойства плазмы в непосредственной окрестности от поверхности обрабатываемого материала изменяются существенно. Проведены исследования потенциала плазмы относительно образца с помощью одиночного электростатического зонда (зонд Ленгмюра) (рис. 5). Результаты исследований показали, что при удалении зонда от поверхности образца положительный потенциал увеличивается примерно по экспоненциальному закону до расстояния 1 мм, а затем потенциал переходит в насыщение (рис. 5а).

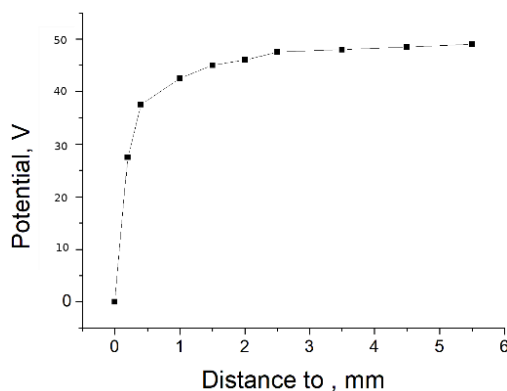


Рис. 5а. Пространственное распределение плавающего потенциала плазмы вблизи обрабатываемого тела ВЧИ-разряда
 $W_p=2,0$ кВт, $P=60$ Па, $G=0,04$ г/с

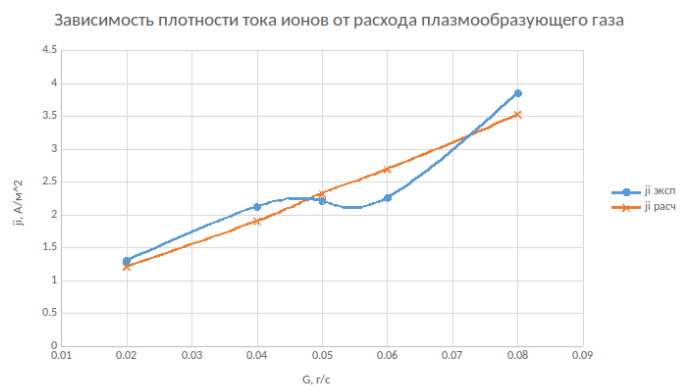


Рис. 5б. Зависимость плотности тока на поверхность от расхода плазмообразующего газа ($p=53,2$ Па, $Pp=1$ кВт)

С увеличением давления в вакуумной камере ВЧ-плазменной установки положительный потенциал плазменного столба монотонно растет, при увеличении расхода газа потенциал плазменного столба уменьшается. При увеличении мощности генератора потенциал плазмы относительно образца монотонно возрастает (рис. 5б).

Сопоставление полученных данных с результатами зондовых измерений позволяет сделать вывод о том, что в окрестности обрабатываемого образца образуется пространственный

слой некомпенсированного положительного заряда толщиной 0,3-0,5 см в зависимости от режима поддержания ВЧИ-разряда пониженного давления. Таким образом, обрабатываемое изделие представляет собой дополнительный электрод, и, следовательно, струя ВЧИ-разряда пониженного давления является не потоком свободно рекомбинирующей плазмы, а представляет собой разряд.

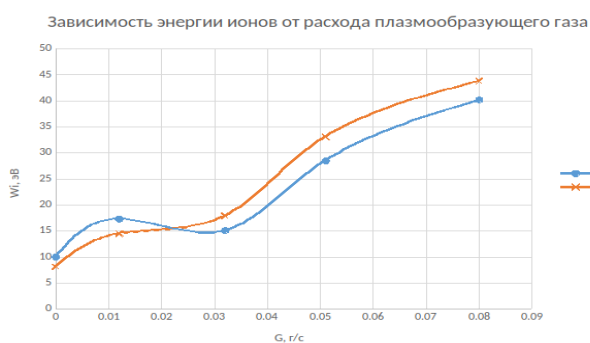


Рис. 6а. Зависимость энергии ионов от расхода плазмообразующего газа ($p = 53,2$ Па, $P = 1$ кВт)

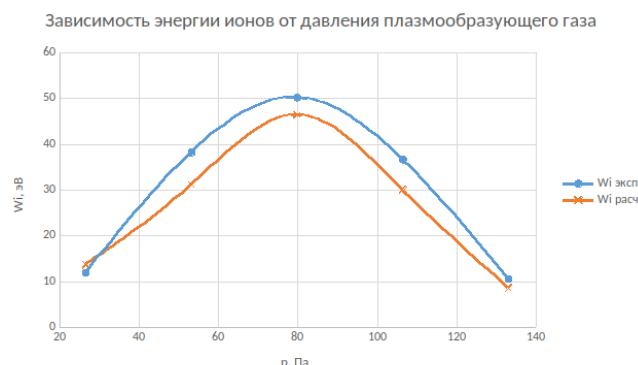


Рис. 6б. Зависимость энергии ионов от давления плазмообразующего газа ($W_p = 1$ кВт, $G = 0,04$ г/с)

Значение максимума энергетического распределения ионов коррелирует с величиной постоянного потенциала плазмы. Зависимости максимумов распределения ионов по энергиям и плотности тока от давления и расхода плазмообразующего газа и мощности ВЧИ-разряда показаны на рис 5б, 6а, 6б. Плотности ионного тока на поверхность образца находились в диапазоне 0,4 – 4,0 А/м² для мощностей разряда 0,6-1,0 кВт. Отклонения расчетных и экспериментальных данных составляют не более 20%.

Зависимость энергии ионов от расхода в ВЧИ-индукционной плазме носят сложный, немонотонный характер. Энергия ионов в плазме аргона на 3-8 эВ больше, чем энергия ионов в смесях аргона с реагирующими газами.

Воздействие низкотемпературной плазмы на поверхность материалов осуществляется в результате ряда сложных, взаимосвязанных процессов энергетического, массового и зарядового обменов частиц плазмы с атомами обрабатываемого тела. Результатом таких взаимодействий являются десорбция атомов и молекул с поверхности тела, распыление и испарение частиц материала, изменения структуры и фазового состояния приповерхностного слоя тела.

При плазменной обработке происходит физическое и химическое взаимодействие материалов с активными и неактивными частицами плазмы, имеющими высокую кинетическую или потенциальную энергию. В процессе обработки полностью разделить физическое и химическое взаимодействия, указать какой-либо один процесс, отвечающий за эффект плазменного воздействия, невозможно.

Процесс обработки включает в себя несколько различных механизмов. Результат обработки, как правило, обусловлен одновременным воздействием на материал различных факторов и определяется параметрами создаваемой плазмы. В промышленных процессах плазменной модификации, в зависимости от свойств низкотемпературной плазмы и зарядового состояния поверхности обрабатываемого материала, можно выделить преимущественный механизм взаимодействия и вид частиц, вносящих наиболее существенный вклад в модификацию поверхности. При бомбардировке твердого тела ионами происходит распыление поверхностного слоя и образование захороненных слоев. В результате снижается шероховатость поверхности и изменяется структура материала в приповерхностном слое на глубине до 150 мкм. Таким образом, наиболее вероятными процессами, ответственными за модификацию поверхности твердых тел в ВЧИ-плазме пониженного давления, являются бомбардировка ионами

низкой (менее 55 эВ) энергии, рекомбинация ионов на поверхности и термическое воздействие.

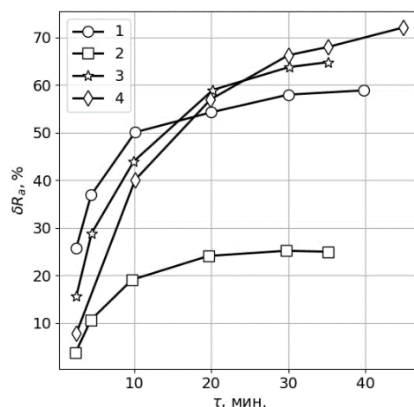


Рис. 7. Зависимость относительного уменьшения параметра шероховатости поверхности образца из ВТ9 от продолжительности обработки ($p = 70$ Па, $G\Gamma = 0,08$ г/с). ВЧИ разряд, $Pp = 2,1$ кВт, $z = 65$ мм, $f = 1,76$ МГц: 1 – аргон, 2 – аргон с 5% азота, 3 – неон, 4 – гелий.

На рис. 7 показана экспериментальная зависимость относительного уменьшения параметра шероховатости поверхности образца из ВТ9 от продолжительности обработки в струйном ВЧИ-разряде пониженного давления. Видно, что со временем процесс выходит на стадию насыщения.

Выводы и заключение. Результаты экспериментальных исследований показали, что ВЧИ-разряд при пониженном давлении с продувом газа существенно отличается от других видов газовых разрядов, в том числе от высокочастотных индукционных разрядов атмосферного давления.

В плазменной струе регистрируются достаточно высокие значения напряженности магнитного поля. При этом наряду с аксиальной составляющей вектора напряженности магнитного поля H_z , обнаруживается азимутальная составляющая H_ϕ . Подтверждено наличие в струе как азимутальной составляющей магнитного поля H_ϕ , так и аксиальной составляющей электрического поля E_z .

Введение твердого тела в струю ВЧИ-разряда пониженного давления существенно изменяет параметры плазмы у поверхности тела, но практически не влияет на разряд в разрядной трубке. При этом:

- у поверхности образца в плазменной струе возникает слой положительного заряда (СПЗ) толщиной до 5 мм, аналогичный приэлектродным СПЗ высокочастотного емкостного разряда. Потенциал СПЗ относительно плазменной струи составляет 30–50 эВ;
- аксиальная составляющая напряженности магнитного поля H_z у поверхности образца уменьшается на 30 – 50%, тогда как азимутальная составляющая H_ϕ возрастает на 20 – 30%;
- энергия ионов, бомбардирующих поверхность образца, составляет 30 – 50 эВ, плотность ионного тока на поверхность 1,5 – 25 А/м²;
- изменения характеристик плазменной струи проявляются одинаково при обработке для металлов, диэлектриков и полупроводников;
- экспериментальные исследования и расчеты [16] подтверждают, что ВЧ-плазменная струя не является потоком рекомбинационной плазмы, так как концентрация электронов на 2 порядка выше в струе, чем в рекомбинирующей плазме.

Впервые, в результате численных экспериментов теоретически предсказан и экспериментально подтвержден эффект образования зоны перегрева по периферии струи около входного отверстия вакуумной камеры, в которой температура потока выше температуры в центре

потока более 50 С при скорости потока на входе в вакуумную камеру более 440 м/с и мощности разряда более 2,4 кВт при расходе плазмообразующего газа более 0,16 г/с. Данный эффект количественно согласуется с результатами экспериментальных исследований.

Приведенные выше экспериментальные данные в сопоставлении с численными расчетами [11-16] доказывают, что струйный ВЧИ-разряд пониженного давления в диапазоне давлений 13,3-133 Па является отдельной разновидностью высокочастотного разряда пониженного давления комбинированного типа, который отличается как от ВЧИ-разряда в потоке газа атмосферного давления с соленоидальным индуктором, так и от ВЧИ-разряда низкого давления в плазмотроне с плоской спиральной антенной, и от ВЧИ-разряда низкого давления с соленоидальным индуктором.

Список источников

1. Абдуллин И.Ш. Модификации нанослоев в высокочастотной плазме пониженного давления /И.Ш. Абдуллин, В.С. Желтухин, И.Р. Сагбиев, М.Ф. Шаехов. – Казань: Изд-во Казан. гос. технол. ун-та, 2007. – 355 с.
2. Абдуллин И.Ш. Высокочастотная плазменно-струйная обработка материалов при пониженных давлениях. Теория и практика применения / И.Ш. Абдуллин, В.С. Желтухин, Н.Ф. Кашапов. – Казань: Изд-во Казан. ун-та, 2000. – 348 с.
3. Bogaerts A., Neyts E., Gijbels R., Van der Mullen J. Gas discharge plasmas and their applications. Spectrochimica Acta Part B: Atomic spectroscopy, 2002, vol. 57, no.4, pp. 609-658.
4. Samukawa S., Hori M., Rauf S. et al. The 2012 plasma roadmap. Journal of physics D, Applied physics, 2012, vol.45, no.25, pp. 253001.
5. Chu P.K., Lu X-P. Low temperature plasma technology: methods and applications. Boca Raton, CRC press, 2013. 315 p.
6. Weltmann K., Kolb J.F., Holub M. et al. The future for plasma science and technology. Plasma processes and polymers, 2019, vol.16, is.1, pp. 1800118.
7. Adamovich I., Baalrud S.D., Bogaerts A. et al. The 2017 plasma roadmap: Low temperature plasma science and technology. Journal of Physics D, Applied physics, 2017, vol.50, no. 32, pp. 323001.
8. Adamovich I., Agarwal S., Ahedo E. et al. The 2022 plasma roadmap: low temperature plasma science and technology. Journal of Physics D, Applied physics, 2022, vol.55, no.37, pp 373001.
9. Райзер Ю.П. Физика газового разряда / Ю.П. Райзер. – М: Наука. Гл. ред. физ.-мат. лит., 1987. 592 с.
10. Boulos M.I. The inductively coupled R.F. (radio frequency) plasma. Pure & applied chemistry, 1985, vol.57, no.9, pp. 1321-1352.
11. Shemakhin A.Y., Zheltukhin V.S., Khubatkuzin A.A. Numerical and experimental study of a warming up effect of an underexpanded rarefied rf plasma jet outflowing into a flooded area. Journal of physics, conference series, 2016, vol. 774, no.1, pp. 012167.
12. Zheltukhin V.S., Shemakhin A.Y. Simulation of rarefied low pressure RF plasma flow around the sample. Journal of physics, conference series, 2017, vol. 789, no. 1, pp. 012071.
13. Абдуллин И.Ш. Сквозная математическая модель струйного высокочастотного индукционного разряда пониженного давления / И.Ш. Абдуллин, В.С. Желтухин, А.Ю. Шемахин // Инженерно-физический журнал, 2024. – Т. 97. – №. 2. – С. 491-503.
14. Шемахин А.Ю. Математическое моделирование ВЧ разряда пониженного давления с продувом газа в нелокальном приближении: электро- и плазмодинамика / А.Ю. Шемахин, В.С. Желтухин, Е.Ю. Шемахин // Инженерно-физический журнал, 2021. – Т. 94. – №. 5. – С. 1368-1376.
15. Shemakhin A. PIGsolver software for simulation RF plasma flow at intermediate pressure. Proc. of 2024 Ivannikov Memorial Workshop (IVMEM). Velikiy Novgorod, IEEE, 2024. pp.49-54.
16. Shemakhin A. Yu. Study of the electron velocity distribution function in weakly ionized radiofrequency plasma. High energy chemistry, 2024, vol.58, no. 1, pp. 147-152.

Шемахин Александр Юрьевич. Кандидат физико-математических наук, доцент кафедры радиофизики института физики Казанского (Приволжского) федерального университета, AuthorID: 644906, SPIN: 4461-4455, ORCID: 0000-0002-4618-3469, ashemakhin@yandex.ru, 420008, г. Казань, ул. Кремлевская 18.

Абдуллин Ильдар Шаукатович. Доктор технических наук, НПО «Плазма ВСТ», AuthorID: 41737, SPIN: 4036-6857, Scopus ID: 7006206963, how80@bk.ru, 420081, г. Казань, ул. Курская, д. 27, офис 1000.

Желтухин Виктор Семенович. Доктор физико-математических наук, профессор, д.н. (доцент) института искусственного интеллекта, робототехники и системной инженерии Казанского (Приволжского) федерального университета, AuthorID: 17329, SPIN: 8945-9662, ORCID: 0000-0003-4452-2110, Vzheldtukhin@gmail.com, 420008, г. Казань, ул. Кремлевская 18.

Макеева Елена Александровна. Аспирант кафедры радиофизики института физики Казанского (Приволжского) федерального университета, ORCID: 0009-0000-0614-664X, elenamakeeva.mackeeva@yandex.ru, 420008, г. Казань, ул. Кремлевская 18.

UDC 537.52:519.624

DOI:10.25729/ESI.2024.35.3.015

Characteristics of a intermediate pressure radio frequency inductive coupled stream discharge for solid processing

Aleksandr Yu. Shemakhin¹, Ildar Sh. Abdullin², Viktor S. Zheltukhin¹, Elena A. Makeeva¹

¹Kazan (Volga Region) Federal University, Russia, Kazan, *ashemakhin@yandex.ru*

²Limited Corporation “Plazma-VST”

Abstract. Experimental measurements of parameters of a stream radio frequency inductive coupled discharge of intermediate pressure in a discharge chamber and a jet, measurements of ion energy and ion current density on the sample surface, validation of a mathematical model of a stream RF discharge of intermediate pressure on experimental data were carried out. It has been established that a stream RF discharge of intermediate pressure in the pressure range of 13.3-133 Pa is a new type of radio frequency intermediate pressure discharge of a combined type, which differs both from an RF discharge in an atmospheric pressure gas stream with a solenoid inductor, and from a low-pressure RF discharge in a plasma torch with a flat spiral antenna, and from an RF discharge-low pressure discharge with a solenoid inductor.

Keywords: stream RF discharge, intermediate pressure, plasma flow, experimental studies, model validation, ICRF

References

1. Abdullin I.Sh., Zheltukhin V.S., Sagbiev I.R., Shaekhov M. F. Modifikacii nanosloev v vysokochastotnoj plazme ponizhennogo davleniya [Modification of nanolayers in a high-frequency plasma of low pressure]. Kazan, 2007, Izd. Kazansk. Gos. Univ., 355 p.
2. Abdullin I.Sh., Zheltukhin V.S., Kashapov N.F. Vysokochastotnaya plazmenno-strujnaya obrabotka materialov pri ponizhennyh davleniyah. Teoriya i praktika primeneniya [High-frequency plasma-jet treatment of materials at low pressures. Theory and practice of application]. Kazan, 2000, Izd. Kazansk. Gos. Univ., 348 p.
3. Bogaerts A., Neyts E., Gijbels R., Van der Mullen J. Gas discharge plasmas and their applications. Spectrochimica Acta Part B. Atomic spectroscopy, 2002, vol. 57, no.4, pp. 609-658.
4. Samukawa S., Hori M., Rauf S. et al. The 2012 plasma roadmap. Journal of physics D, Applied physics, 2012, vol.45, no.25, pp. 253001.
5. Chu P.K., Lu X-P. Low temperature plasma technology: methods and applications. Boca Raton, CRC press, 2013. 315 p.
6. Weltmann K., Kolb J.F., Holub M. et al. The future for plasma science and technology. Plasma processes and polymers, 2019, vol.16, is.1, pp. 1800118.
7. Adamovich I., Baalrud S.D., Bogaerts A. et al. The 2017 plasma roadmap: Low temperature plasma science and technology. Journal of Physics D, Applied physics, 2017, vol.50, no. 32, pp. 323001.
8. Adamovich I., Agarwal S., Ahedo E. et al. The 2022 plasma roadmap: low temperature plasma science and technology. Journal of Physics D, Applied physics, 2022, vol.55, no.37, pp 373001.
9. Rajzer Yu. P. Fizika gazovogo razryada [Physics of gas discharge]. Moscow, 1987, Nauka [Science], 592 p.
10. Boulous M.I. The inductively coupled R.F. (radio frequency) plasma. Pure & applied chemistry, 1985, vol.57, no.9, pp. 1321-1352.
11. Shemakhin A.Y., Zheltukhin V.S., Khubatkhuzin A.A. Numerical and experimental study of a warming up effect of an underexpanded rarefied rf plasma jet outflowing into a flooded area. Journal of physics, conference series, 2016, vol. 774, no.1, pp. 012167.

12. Zheltukhin V.S., Shemakhin A.Y. Simulation of rarefied low pressure RF plasma flow around the sample. Journal of physics, conference series, 2017, vol. 789, no. 1, pp. 012071.
13. Abdullin I.Sh., Zheltukhin V.S., Shemakhin A.Y. Skvoznaya matematicheskaya model' strujnogo vysokochastotnogo indukcionnogo razryada ponizhennogo davleniya [End-to-end mathematical model of a radio-frequency inductive jet discharge of lowered pressure]. Inzhenerno-fizicheskij zhurnal [Journal of engineering physics and thermophysics], 2024, vol.97, no.2, pp. 491-503.
14. Shemakhin A.Y., Zheltukhin V.S., Shemakhin E.Y. Matematicheskoe modelirovanie VCh razryada ponizhennogo davleniya s produvom gaza v nelokal'nom priblizhenii: elektro- i plazmodinamika [Mathematical simulation of a high-frequency low-pressure discharge with gas injection in non-local approximation: electro- and plasmadynamics]. Inzhenerno-fizicheskij zhurnal [Journal of engineering physics and thermophysics], 2021, vol.94, no. 5, pp. 1368-1376.
15. Shemakhin A. PIGsolver software for simulation RF plasma flow at intermediate pressure. Proc. of 2024 Ivannikov Memorial Workshop (IVMEM). Velikiy Novgorod, IEEE, 2024. pp.49-54.
16. Shemakhin A. Yu. Study of the electron velocity distribution function in weakly ionized radiofrequency plasma. High energy chemistry, 2024, vol.58, no. 1, pp. 147-152.

Shemakhin Aleksandr Yuryevich. *Cand. sci., associate professor of department of radiophysics of institute of physics, Kazan (Volga Region) federal university, AuthorID: 644906, SPIN: 4461-4455, ORCID: 0000-0002-4618-3469, ashemakhin@yandex.ru, 420008, Kazan, Kremlevskaya str., 18.*

Abdullin Ildar Shaukatovich. *Doctor of technical sciences, limited corporation "Plazma-VST", AuthorID: 41737, SPIN: 4036-6857, Scopus ID: 7006206963, how80@bk.ru, 420081, Kazan, Kurskaya str., 27, office 1000.*

Zheltukhin Viktor Semenovich. *D.Sc., professor, associate professor of institute of artificial intelligence, robotics and system engineering, Kazan (Volga Region) federal university, AuthorID: 17329, SPIN: 8945-9662, ORCID: 0000-0003-4452-2110, Vzheltukhin@gmail.com, 420008, Kazan, Kremlevskaya str., 18.*

Makeeva Elena Alexandrovna. *Postgraduate student, department of radiophysics of institute of physics, Kazan (Volga Region) federal university, ORCID: 0009-0000-0614-664X, elenamakeeva.mackeeva@yandex.ru, 420008, Kazan, Kremlevskaya str., 18.*

Статья поступила в редакцию 30.09.2024; одобрена после рецензирования 11.10.2024; принята к публикации 14.10.2024.

The article was submitted 09/30/2024; approved after reviewing 10/11/2024; accepted for publication 10/14/2024.

**Основные рубрики журнала
«Информационные и математические технологии
в науке и управлении»**

№ Основная рубрика	Составляющие (используемые ранее рубрики)
1. Методологические аспекты информационных и математических технологий (1.2.2, 2.3.1, 2.3.3)	Математические и информационные технологии Методы и средства обеспечения кибербезопасности Информационные и математические технологии Информационно-вычислительные технологии Вычислительные технологии Математические технологии
2. Методы, технологии и приложения искусственного интеллекта (1.2.2, 2.3.1, 2.3.5)	Методы и модели искусственного интеллекта Методы искусственного интеллекта Методы искусственного интеллекта и их приложения Методы и приложения искусственного интеллекта Методы и системы искусственного интеллекта Искусственный интеллект Искусственный интеллект, онтологии Технологии искусственного интеллекта Технологии искусственного интеллекта и их применение Применение методов искусственного интеллекта Семантическое моделирование Экосистема знаний, базы знаний
3. Математическое моделирование и его применение в научных исследованиях (1.2.2, 2.3.5, 2.3.3, 2.3.1)	Методы математического моделирования Математическое моделирование Математическое моделирование в технике Математическое моделирование в технике и экономике
4. Математические, информационные и интеллектуальные технологии в энергетике (2.4.5, 1.2.2, 2.3.1, 2.3.3)	Математическое моделирование в энергетике Математические технологии в энергетике Информационные и математические технологии в энергетике Интеллектуальные технологии в энергетике Математическое моделирование и кибербезопасность в энергетике

5. Цифровая экономика и управление (1.2.2, 2.3.1., 2.4.5, 2.3.3)	Цифровые технологии в энергетике и экономике Математическое моделирование в энергетике и экономике Методы и средства обеспечения кибербезопасности Цифровые двойники Цифровые платформы Инновации и информационные технологии Поддержка принятия решений Системы поддержки принятия решений
6. Информационные системы и технологии (1.2.2, 2.3.1, 2.3.5)	Информационные технологии и программные системы Информационные и интеллектуальные технологии Информационные и математические технологии Информационные технологии Информационные технологии в образовании Информационно-вычислительные технологии Информационные системы и программные комплексы Информационные и программные системы Базы данных
7. Программные системы и комплексы (2.3.5, 1.2.2, 2.3.3, 2.3.1)	Программные комплексы и информационные системы Программные системы Информационные системы и программные комплексы Информационные и программные системы Цифровые двойники Визуальная аналитика

Для справки: Специальности ВАК РФ

- **1.2.2.** Математическое моделирование, численные методы и комплексы программ (технические науки) (бывшая 05.13.18),
- **2.3.1.** Системный анализ, управление и обработка информации (технические науки) (05.13.01),
- **2.3.3.** Автоматизация и управление технологическими процессами и производствами (технические науки) (05.13.06),
- **2.3.5.** Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки) (05.13.11),
- **2.4.5.** Энергетические системы и комплексы (технические науки) (05.14.01).

