

НГТУ
НЭТИ

—

C_H CHO
НЭТИ

—

СН

НОВОСИБИРСК
2025



ВСЕРОССИЙСКАЯ НАУЧНАЯ
КОНФЕРЕНЦИЯ МОЛОДЫХ УЧЁНЫХ

НАУКА. ТЕХНОЛОГИИ.
ИННОВАЦИИ

СБОРНИК НАУЧНЫХ ТРУДОВ 01

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

НАУКА ТЕХНОЛОГИИ ИННОВАЦИИ

Сборник
научных трудов

г. Новосибирск, 02 декабря – 06 декабря 2024 г.
в восьми частях

Часть 1

НОВОСИБИРСК
2025

УДК 62(063)
ББК 72.5я431
НЗ4

НЗ4 НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ // Сборник научных трудов
в 8 ч. / под редакцией Захаровой Е. В. – Новосибирск: Изд-во НГТУ, 2025.

ISBN 978-5-7782-5342-1
Часть 1: – 412 с.
ISBN 978-5-7782-5343-8

В сборнике публикуются материалы по научным направлениям:
*«Информатика, автоматика, вычислительная и измерительная техника»,
«Информационные технологии математического моделирования и обработки
данных».*

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ

Председатель Оргкомитета:

Батаев А. А., д.т.н., профессор, ректор НГТУ

Зам. председателя Оргкомитета:

Отто А. И., к.т.н., проректор по научной работе и инновациям НГТУ

Программный комитет:

Драгунов В. П., д.т.н., профессор, начальник ОПК ВК

Корель И. И., к.ф.-м.н., доцент, декан ФТФ

Осьмук Л. А., д.соц.н., профессор, директор ИСТР

Рева И. Л., к.н.т., доцент, декан АВТФ

Мелехина Е. А., к.п.н., доцент, декан ФГО

Чинахов Д. А., д.т.н., доцент, декан ФЛА

Тимофеев В. С., д.т.н., доцент, декан ФПМИ

Эзрох Ю. С., д.э.н., доцент, декан ФБ

Стрельцов С. А., к.т.н., доцент, декан РЭФ

Вильбергер М. Е., к.т.н., доцент, декан ФМА

Тюрин А. Г., к.т.н., доцент, декан МТФ

Русина А. Г., д.т.н., профессор, декан ФЭН

**УДК 62(063)
ББК 72.5я431**

**ISBN 978-5-7782-5343-8 (Ч.1)
ISBN 978-5-7782-5342-1**

© Коллектив авторов, 2025
© Новосибирский государственный
технический университет, 2025

**НАУЧНОЕ НАПРАВЛЕНИЕ – ИНФОРМАТИКА,
АВТОМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ
И ИЗМЕРИТЕЛЬНАЯ ТЕХНИКА**

**Секция АВТОМАТИКА, ИЗМЕРЕНИЯ
И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ**

**ИССЛЕДОВАНИЕ ПОДАВЛЕНИЯ ПУЛЬСАЦИЙ И ШУМОВ
DC/DC-ПРЕОБРАЗОВАТЕЛЯ MORN SUN B2415S-2WR3**

К.А. Волобуев, Д.Н. Катасонов, В.Г. Трубин
Новосибирский государственный технический университет,
г. Новосибирск, kirya.volobuev@mail.ru, katasonov@corp.nstu.ru,
trubin@corp.nstu.ru

Научный руководитель: Катасонов Д.Н., к.т.н.

В данной работе проведено исследование подавления пульсаций и шумов на выходе DC/DC-преобразователя. Произведено сравнение нескольких методов фильтрации и приведены рекомендации по их применению.

This paper studies the suppression of pulsations and noise at the output of a DC/DC converter. Several filtration methods are compared and provided recommendations for their use.

Введение

При разработке электронных устройств, может потребоваться дополнительное питающее напряжение или гальваническая развязка питания. Эти задачи можно решить использованием импульсных DC/DC-преобразователей (конверторов). Основными недостатками преобразователей данного типа являются пульсации и шумы на выходе, борьба с которыми может значительно усложнить схемотехнику устройства.

Известным производителем источников питания является китайская компания Mornsun [1], которая недавно выпустила на рынок третье поколение (R3) изолированных DC/DC-преобразователей. В данной работе проведено исследование уровня пульсаций и шумов на выходе конвертора B2415S-2WR3, а также рассмотрены методы их подавления.

Исследование выхода преобразователя

Для корректного тестирования преобразователя необходимо исключить длинные провода, так как они будут выступать как антенны и ловить наводки, амплитуда которых может превышать амплитуду

пульсаций и шумов преобразователя. Таким образом, сборка на контактной макетной плате недопустима. Поэтому, схема, представленная на рисунке 1, была смонтирована на печатной макетной плате. В качестве нагрузки преобразователя выступает светодиодная лента, чтобы не превысить максимальный ток конвертера установлен токоограничительный резистор (ток нагрузки 120 мА). На рисунке 2 приведена осциллограмма с выхода (клемма P1).

Пульсации выходного напряжения в DC/DC-преобразователях возникают из-за циклов зарядки и разрядки выходного конденсатора. На пульсации накладываются шумы (короткие выбросы напряжения с высокой частотой), вызванные работой ключей генератора.

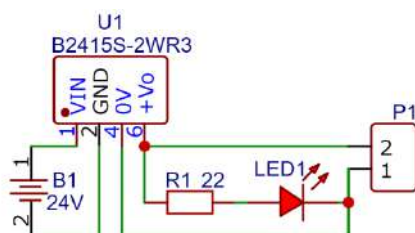


Рис. 1 – Схема подключения DC/DC-преобразователя

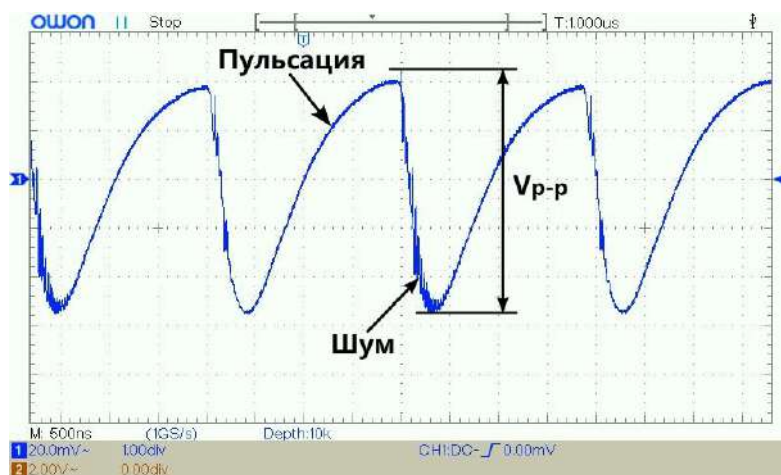


Рис. 2 – Пульсации и шумы на выходе DC/DC-преобразователя

Подавление пульсаций преобразователя

Подавление пульсаций и шумов необходимо производить с помощью дифференциальной и синфазной фильтратий.

Простейшую дифференциальную фильтрацию, можно осуществить, установив на выходе внешний конденсатор C_{EXT} , желательно с низким значением параметра ESR. Однако при подключении большой ёмкости могут возникнуть проблемы запуска DC/DC-преобразователя.

Более рациональным способом уменьшения пульсаций выходного напряжения будет использование LC-фильтра низких частот (ФНЧ). В случае добавления выходной индуктивности L_{EXT} рассчитать уровень выходных пульсаций V_{p-p} можно по формуле:

$$V_{p-p} = \frac{I_{OUT}}{2f_{OPER}(C_{OUT} + \sqrt{L_{EXT}C_{EXT}})},$$

где C_{OUT} – ёмкость встроенного конденсатора преобразователя. Но схема и номиналы компонентов DC/DC-преобразователя не всегда известны. В таком случае, чтобы достичь эффективного снижения пульсаций выходного напряжения рекомендуется частоту среза LC-фильтра принять равной 1/10 от значения рабочей частоты преобразователя f_{OPER} .

$$f_c = f_{OPER}/10 = \frac{1}{2\pi} \sqrt{\frac{1}{LC}}$$

На частоте среза фильтра f_c сигнал ослабляется на –3 дБ. ФНЧ на основе LC-цепи имеет второй порядок (изменение сигнала –40 дБ/дек), таким образом, сигнал с частотой, которая в 10 раз выше, чем частота среза фильтра, будет уменьшен в 100 раз.

Чтобы произвести фильтрацию синфазных шумов требуется установить синфазный дроссель, после которого также рекомендуется поставить выходной конденсатор.

На рисунке 3 приведена схема тестовой установки. Как и ожидалось, наилучший результат (рисунок 4) был получен при замкнутых ключах SW2, SW5 и разомкнутых SW1, SW3, SW4.

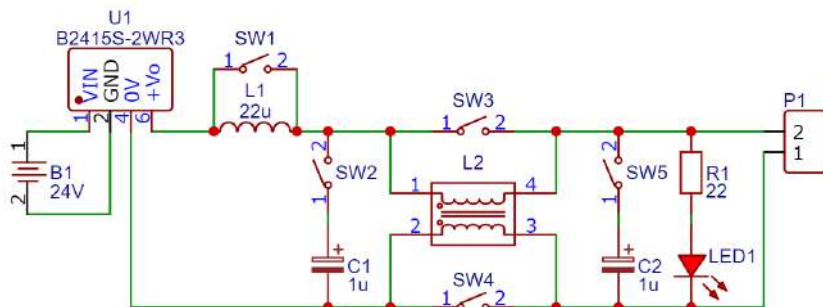


Рис. 3 – Схема подключения DC/DC-преобразователя для исследования фильтрации

В таблице 1 приведены результаты для разных положений ключей, где 1 – ключ замкнут, 0 – разомкнут.

Таблица 1 – Результаты измерений

SW1	SW2	SW3	SW4	SW5	V_{p-p} , мВ
1	0	1	1	0	100
1	0	1	1	1	69
0	1	1	1	1	20
0	1	0	0	1	12

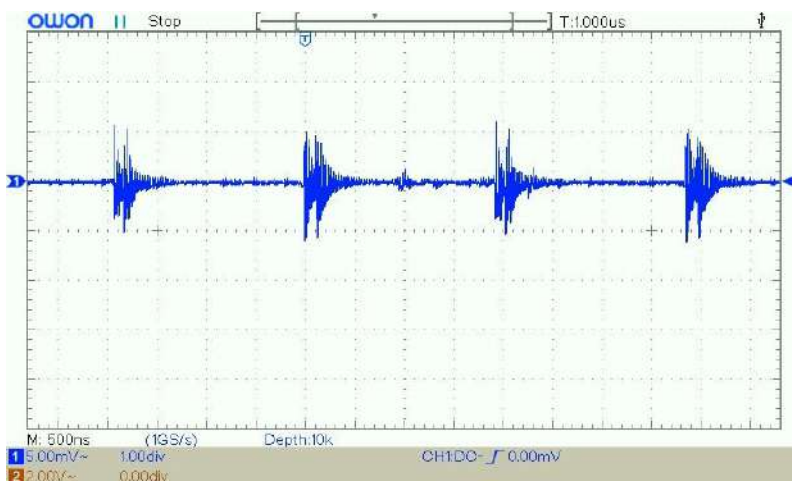


Рис. 4 – Пульсации и шумы на выходе DC/DC-преобразователя при замкнутых ключах SW2, SW5 и разомкнутых SW1, SW3, SW4.

Выводы

Осциллограмма, приведённая на рисунке 4, показывает, что с использованием дополнительных фильтров пульсации на выходе схемы практически исчезли, остались только высокочастотные шумы, но их амплитуда в разы меньше, чем была изначально. В схеме без использования фильтров напряжение от пика до пика около 100 мВ, после установки фильтров оно составляет около 12 мВ.

Данное исследование проводилось при питании DC/DC-преобразователя от аккумуляторов. В реальных системах рекомендуется устанавливать по входу фильтры, аналогичные приведённым выходным фильтрам. Это не только отфильтрует входное напряжение, приходящее на преобразователь, но и подавит обратные пульсации и шумы, которые генерируют ключи преобразователя. Следует отметить, что на амплитуду пульсаций и шумов может влиять расположение компонентов и дорожек на плате.

Литература:

1. Документация на DC/DC-преобразователи Mornsun поколения 3. – URL: <https://www.mornsun-power.com/html/pdf/B2415S-2WR3.html> (дата обращения: 20.10.2024).
2. *Roberts S.* DC/DC Book of Knowledge Practical tips for the User. RECOM Third Edition [Electronic resource] // RECOM. – 2016. – URL: <https://www.micro-semiconductor.com/datasheet/90-DC-DC-BOOK-OF-KNOWLEDGE-EN.pdf> (дата обращения: 20.10.2024).
3. Wide Input Voltage DC/DC Converters Application Guide 2021. [Electronic resource] // Mornsun. – 2021. – URL: MORNSUN Wide Input Voltage DC DC Converters Application Guide 2021 (дата обращения: 20.10.2024).
4. *Рентюк В.* Как правильно выбрать фильтр для изолированных DC/DC-преобразователей? [Электронный ресурс] // Силовая электроника. – № 3. – 2019. – URL: https://mstator.ru/publications/2019/filter_for_isolated_DC-DC (дата обращения: 20.10.2024).
5. *Roberts S., Stangl A.* Как сделать DC/DC-преобразователи самыми эффективными [Электронный ресурс] // Электронные компоненты. – №10. – 2020. – URL: <https://elcomdesign.ru/wp-content/uploads/2020/10/44.pdf> (дата обращения: 20.10.2024).

РАЗРАБОТКА АЛГОРИТМА РАСПОЗНАВАНИЯ КОНТУРОВ НА ПРОФИЛОГРАММАХ ПРИ ПОМОЩИ СИСТЕМЫ ТЕХНОЛОГИЧЕСКОГО ЗРЕНИЯ

К.А. Цирульникова, Д.Н. Катасонов

**Новосибирский государственный технический университет,
г. Новосибирск, ksenia-7500@yandex.ru**

Научный руководитель: Катасонов Д.Н., к.т.н., доцент

В данной статье рассматриваются различные подходы и методы по поиску контуров на изображениях. Исследования в этой области необходимы для разработки алгоритма, направленного на оптимизацию процесса анализа и проверки качества поверхности, проводимых на основе изображений, получаемых профилометром. Разработка алгоритма производится на языке программирования C++ с применением библиотеки компьютерного зрения OpenCV, а также таких методов как выделение дескрипторов особых точек, детектор границ Канны и цепной код Фримана.

This article discusses various approaches and methods for finding contours in images. Research in this area is necessary to develop an algorithm aimed at optimizing the process of analyzing and verifying surface quality based on images obtained by a profiler. The algorithm is developed in the C++ programming language using the OpenCV computer vision library, as well as methods such as the allocation of special point descriptors, the Canny boundary detector and the Freeman chain code.

Введение. Профилометрия (процесс измерения профиля сечения поверхности в плоскости) представляет собой неотъемлемую часть в различных областях науки и производства. В результате работы профилометра формируется так называемая карта высот или профилограмма – набор расстояний от измеряемой поверхности до измерительного инструмента. Профилограмма обычно формируется в виде изображения, где измеренное расстояние кодируется в условных цветах, на его основе производится оценка параметров, исследуемых для частных случаев, или же на основе стандартизированных параметров шероховатости [1, 5]. Оценка поверхности позволяет отслеживать ее качество, а в условиях производства осуществлять проверку на соответствие заданным допустимым критериям.

Проблематика проверки соответствия поверхности заданным нормам в условиях серийного производства состоит в большом потоке исследуемых объектов. Как правило, на обрабатываемом объекте присутствует тестовый элемент, относительно которого производится проверка на качество выполненной операции, для ее осуществления оператору вручную необходимо найти и выделить тестовый элемент,

данная операция требует дополнительных временных затрат. При этом, профилометр может иметь систему технического зрения, которая может быть использована для поиска интересующей зоны на измеряемом объекте. Таким образом, возникает две задачи поиска необходимого фрагмента на изображении. В первом случае происходит поиск зоны интереса на изображении, а во втором случае при анализе сформированной профилограммы. В обоих случаях возможно предварительное формирование эталонного изображения.

Целью разрабатываемого алгоритма по обнаружению контура в соответствии с заданным эталоном является автоматизация вышеописанной операции на серийном производстве. После обнаружения необходимого контура возможна дальнейшая оптимизация работы путем добавления автоматического расчета параметров шероховатости или же других, требуемых критериев для оценки качества поверхности.

Методы и материалы. Особенностью разрабатываемого алгоритма является гибридизация различных методов, присутствующих в библиотеке OpenCV, а также особенности предметной области его применения. Решаемую задачу можно разделить на три подзадачи: нормализация исследуемой области; детектирование границ; поиск в соответствии с эталоном.

Нормализация исследуемой области. Нормализация для данного случая заключается в корректировке угла поворота изображения. При этом корректировка масштаба изображения не требуется – на этапе формирования изображения системой технического зрения масштаб задается параметрами объектива и параметрами фотоматрицы, а при формировании профилограммы исходя из заранее задаваемого оператором шага сканирования (размера пиксела). Угол поворота, в свою очередь, может варьироваться из-за влияния внешних факторов, возникновение которых допускается при размещении образца на платформе профилометра.

Для осуществления операции рассматривалось два подхода: вычисление дескрипторов ключевых точек и итерационный алгоритм совмещения контуров с неравномерным шагом дискретизации [2]. При сравнении обоих подходов была принята во внимание скорость их выполнения, лучший результат показал подход вычисления особых точек. Среди подобных алгоритмов был выбран ORB, как один из самых быстрых и эффективных, сочетающий в себе идеи алгоритмов FAST и BRIEF.

Детектирование границ. Для выделения контуров выбран детектор границ Канны. Его преимуществом является невосприимчивость к шуму

на изображении, в отличие от таких алгоритмов, как оператор Робертса или Собеля. Алгоритм Канни состоит из следующих шагов: размытие исходного изображения с помощью функции Гаусса; поиск градиента, по максимальному значению которого выделяются границы; подавление не-максимумов, для которого устанавливаются пороговые значения [3].

Исходя из постановки задачи, где сказано, что в обязательном порядке присутствует тестовый элемент с уникальной формой, мы можем производить сравнение, оценивая формы контуров. Для реализации данной идеи подойдет цепной код Фримена, идея которого заключается в представлении границы в виде последовательности отрезков прямых линий установленной длины и направления, которые кодируются значениями, исходя из правила [4]. Таким образом, последовательно обходя контур, формируется его цепной код, на рисунке 1 показано правило кодирования и пример.

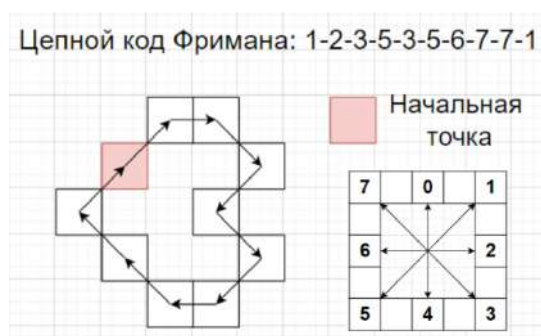


Рис.1 - Соответствие направлений отрезков их кодировке и пример построения цепного кода

Поиск в соответствии с эталоном. Так как цепной код разных контуров имеет разную длину, для корректного сравнения используется гистограмма направлений, формирующаяся через анализ частоты вхождения: подсчитывается количество вхождений в цепной код каждого направления, представленного кодом от 0 до 7. Для сравнения гистограмм эталонного и рассматриваемого контуров применяется косинусное сходство, которое измеряет угол между двумя гистограммами, представленными в виде вектора. По результатам вычисления косинуса найденного угла производится оценка - чем ближе значение косинуса к 1 (угол ближе к **Ошибка! Ожидалась цифра.**), тем больше схожи рассматриваемые гистограммы.

Результаты. Работа описанного алгоритма, состоящего в гибридизации методов, подобранных на основе теоретических данных или практических опытов, обеспечивает обнаружение заданного контура, а также отображение соответствия эталонного контура к найденному в процентах. Тестирование производилось на образцах из пластика, стекла, на микросхемах. Пример работы на рисунках 2-6:

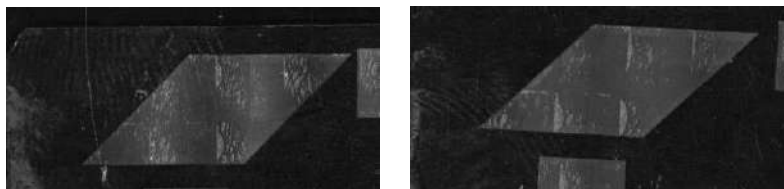


Рис.2-3 – Изображения образца из стекла, полученные камерой профилометра (слева – эталонное, справа – исследуемое)



Рис.4-5 – Результат работы детектора границ Канни после нормализации – слева; границы, представленные цепным кодом Фримана – справа



Рис.6 – Выделенный найденный контур; соответствие – 91%

Заключение. Главная идея данной статьи заключается в разработке алгоритма, примененного к такой области решения задач, как поиск контуров на изображениях, получаемых с помощью профилометра, в условиях серийного производства, где существует необходимость в оптимизации работы оператора. Направлениями дальнейших исследований могут являться сокращение времени, необходимого для выполнения и повышение автоматизации получения информации о свойствах объектах при помощи профилометра.

Литература:

1. *Кравчук М.А., Крайко С.Э., Шелег В.К.* Определение шероховатости обработанной поверхности: методические указания – Минск: БНТУ, 2019. – 25 с.

2. *Дязитдинов Р.Р.* Итерационный алгоритм совмещения контуров с неравномерным шагом дискретизации [Электронный ресурс] // Компьютерная оптика - Т.47, № 1. – 2023. – URL: <https://www.computeroptics.ru/> (дата обращения: 05.11.2024). – doi: 10.18287/2412-6179-CO-1123.

3. *Asif Fuad A.* Hardware Software CoSimulation of Canny Edge Detection Algorithm. International Journal of Computer Applications (0975 – 8887). [Electronic resource] // vol. 122, №19. – 2015. URL: <http://research.ijcaonline.org/volume122/number19/pxc3905124.pdf/> (accessed 05.11.2024).

4. *Freeman H.* On the encoding of arbitrary geometric configurations [Electronic resource] // IEEE Electronic Computers, IRE Transactions, - vol. 10, No 2. - 1961. URL: <https://ieeexplore.ieee.org/document/5219197> (accessed 05.11.2024) - doi: 10.1109/TEC.1961.5219197.

5. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. – Москва, 2005. – 1072 с.

ОБСУЖДЕНИЕ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ ОПТИЧЕСКОГО ВОЛОКНА ДЛЯ ПЕРЕДАЧИ ИНФОРМАЦИИ В СЛОЖНЫХ ПРОМЫШЛЕННЫХ УСЛОВИЯХ

И.М. Ребус

**Новосибирский государственный технический университет,
г. Новосибирск, ilya.rebus42@gmail.com**

Научный руководитель: Французова Г.А., д.т.н., доцент

В статье обсуждаются особенности использования оптического волокна и медной витой пары для высокоскоростной передачи данных в условиях подземных горнодобывающих предприятий. Показано, что несмотря на эволюцию и распространение оптоволоконных технологий связи применение витой пары в опасных и труднодоступных условиях горных выработок остается на сегодняшний день предпочтительным.

This article researches pros and cons of optic-fiber cable usage for high-speed data transmission in underground mining facilities. The main goal of this article is to justify the continued use of copper twisted-pair cables for high-speed, high-distance data transmission despite evolution and spreading of optic-fiber communication technologies.

Введение. Оптическое волокно представляет собой тонкую нить из оптически прозрачного материала (например, кварцевого стекла), заключённую в оболочку из материала с отличающимся показателем преломления света. При подаче светового потока из оптически более плотной среды в менее плотную, отражение электромагнитной волны происходит без потери энергии импульса. Это делает возможным передачу данных на расстояния от нескольких сотен метров до нескольких сотен километров.

Данная особенность уже давно сделала оптоволокно основным способом передачи данных на большие расстояния. Оно связывает города, страны и континенты, уверенно вытесняя медные провода. Последние остаются широко востребованными, как правило, в качестве решения проблемы «последней мили» - организации соединения конечного абонента с устройством-концентратором оператора линии связи. Зачастую эти расстояния не превышают одной-двух сотен метров.

Современные шахты и рудники оборудуются компьютеризированными системами безопасности и сигнализации, что значительно повышает безопасность труда сотрудников предприятия и помогает предотвратить катастрофы. Поскольку длина горных выработок на рудниках или угольных шахтах может достигать нескольких километров, то проблема высокоскоростной передачи данных под землёй является актуальной. Обсудим возможность применения с этой целью оптоволокна, выделив его достоинства и недостатки [1, 2].

Достоинства оптоволоконных линий связи. Самым очевидным преимуществом является *высокая скорость и дальность* связи. В отличие от медной витой пары, где такие параметры как сбалансированность пары, импеданс, затухание сигнала, задержка влияют в значительной степени, в оптическом кабеле они либо отсутствуют вовсе, либо влияют в значительно меньшей степени.

Оптическое волокно имеет хорошую *помехозащищённость*. Оно не проводит электрический сигнал и является изолированной средой, а следовательно никакие виды электромагнитных помех не влияют на качество передачи информации. Таким образом, его можно прокладывать близко к таким источникам электромагнитных помех в шахте, как силовые электрические кабели, трансформаторы, конвейерные ленты и т. д.

Оптические волокна отличают *малые габариты, масса и стоимость*. Оно тоньше и легче медных жил, что облегчает конструкцию кабеля. Стоимость оптического кабеля на километр

расстояния меньше стоимости УТР-кабеля и проще в транспортировке, а значит снижаются затраты как на покупку самого кабеля, так и на его доставку к месту монтажа.

Недостатки оптоволоконных линий связи. Качество передачи данных напрямую зависит от показателей преломления сердцевины и оболочки и наличия подходящих условий для проявления эффекта полного внутреннего отражения. Деформация кабеля нарушает эти условия, и часть светового потока выходит в оболочку. Выход светового потока приводит к затуханию сигнала, поэтому у любого оптического кабеля имеется минимально допустимый радиус изгиба, при котором гарантируется отсутствие затуханий. Помимо этого, материал, из которого изготавливается волокно, хрупок по своей природе, а следовательно его легко повредить при неосторожном обращении. Что в условиях активных проходческих работ происходит очень часто, как и излишняя деформация и перегибание кабеля. Следовательно, *хрупкость и большой минимальный радиус изгиба оптоволоконна* являются основным недостатком, ограничивающим его применение в подземных выработках.

При монтаже или ремонте оптоволоконного кабеля, его волокна необходимо правильно сварить друг с другом. Это долгий и трудоёмкий процесс, требующий специальных навыков и знаний, а следовательно, квалифицированного работника. При этом требуется специальный сварочный аппарат, стоимость которого колеблется в пределах нескольких сотен тысяч рублей. Таким образом, *сложность монтажа и ремонта* затрудняет применение оптоволоконна в горнодобывающей промышленности.

Подземные горнодобывающие предприятия – опасная производственная среда, в которой любое электронное оборудование должно соответствовать требованиям ГОСТ 30852.0-2002 о взрывозащищённости в условиях взрывоопасных газовых сред [3]. Поэтому даже при наличии специализированного оборудования для сварки оптоволоконна, применять его в самой шахте запрещено, если оно не соответствует требованиям о *взрывозащищённости*. Повреждённый кабель нужно будет сначала поднять на поверхность, а уже потом производить ремонт.

Из вышеперечисленных недостатков следует, что при повреждении оптического кабеля на его ремонт уйдёт много времени и средств, причем в сложных условиях горных выработок повреждения будут происходить довольно часто. При отсутствии надлежащего дублирования, во время ремонта линия связи не будет работать, поэтому оборудование, обеспечивающее мониторинг и безопасность,

окажется выведено из строя, что повлечёт за собой остановку производства и значительные финансовые потери.

Таким образом, оптоволокну может быть использовано на магистральных линиях связи, располагающихся в основных проходах, соединяющих разные части шахты между собой и с поверхностью, где не проводятся активные горные работы.

Применение медной витой пары для передачи информации.

Связью и системами безопасности нужно обеспечивать не только отдельные наземные части горнодобывающих предприятий, но и лавы, где ведётся непосредственно добыча полезных ископаемых, и которые могут иметь довольно большую протяжённость. В этом случае витая пара оказывается предпочтительным выбором, поскольку лишена всех вышеприведённых недостатков оптоволокну. К ее недостаткам следует отнести *низкую скорость и дальность* передачи информации, а также *подверженность помехам*.

Таблица 1 – сравнение способов реализации канала связи

	Витая пара (UTP 5e)	ТПП 10х2х0.5	Опт. кабель ОКСК-4А-10
Количество изолированных проводников	8	20	4
Скорость передачи данных на 1 км (Мбит/с)	10 (BASE-T1L)	10 (BASE-T1L)	От 100 до 100 000 (OM1)
Стоимость за 1 м (руб)	22.92	60	14.15
Температура эксплуатации (°C)	-20..+60	-50..+60	-60..+70
Вес нетто 1 км кабеля (кг)	29.5	96	25
Мин. рад. изгиба (наружных диаметров)	8	10	20
Стоимость спец. оборудования (руб)	-	-	от 200 000
Наличие квалифицированных кадров для обслуживания	Нет	Нет	Да

Заключение. Несмотря на все преимущества и широкое распространение оптоволоконной линии связи, применение ее в горной промышленности ограничено наземным использованием. В условиях подземных работ для передачи данных на большие расстояния целесообразно использование медной витой пары. При этом актуальной задачей является развитие соответствующих технологий, направленных на повышение скорости передачи информации и помехозащищенности линии связи, что предполагает проведение исследований в этой области.

Литература:

1. Семенов А.Б. Волоконно-оптические подсистемы современных ККС. – Изд-во ДМК Пресс, Компания АйТи, 2014. – 632 с.
2. Michael K. Barnoski. Fundamentals of Optical Fiber Communications, Second Edition. – Academic Press, 2012. – 364 p.
3. Олифер Н. А. Компьютерные сети. Принципы, технологии, протоколы / Учебник для вузов. 3-е изд. — СПб.: Питер, 2006.
4. IEEE 802.3cg-2019 standard for Ethernet.
5. ГОСТ 30852.0-2002. Электрооборудование взрывозащищенное.
6. ГОСТ Р 59283-2020. Оборудование горно-шахтное. Многофункциональные системы безопасности угольных шахт. Аэрогазовый контроль. Сканирующий контроль метана и оксида углерода. Общие технические требования.
7. ГОСТ Р 55154-2012. Оборудование горно-шахтное. Системы безопасности угольных шахт многофункциональные. Общие технические требования.

РАЗРАБОТКА ПРОГРАММНО-АППАРАТНОГО КОМПЛЕКСА ДЛЯ ИЗМЕРЕНИЯ ЭЛЕКТРИЧЕСКОГО ПОТЕНЦИАЛА СЕНСОРА ГЛЮКОЗЫ

**А.А. Каштанов, И.Л. Рева, И.А. Фуряев, Е.В. Каштанова,
Новосибирский государственный технический университет,
г. Новосибирск, kashtanov.alex@inbox.ru
Научный руководитель: Рева И.Л., к.т.н., доцент**

В работе рассматривается электронная схема и программное обеспечение потенциостата, предназначенного для измерения уровня глюкозы в растворе. Ключевыми критериями при разработке была низкая стоимость, компактный размер и достаточная точность измерений. Результаты измерений выводятся в виде графиков для возможности производить визуальное сравнение с эталонным измерением.

The paper considers the development of an electrical circuit and software for a potentiostat designed to measure glucose levels in a solution. The key criteria in the development were low cost, compact size, and sufficient measurement accuracy. The measurement results are displayed in the form of graphs to enable visual comparison with a reference measurement.

В текущий момент не существует готовых решений для компактного и недорогого потенциостата, хотя потребность в данном приборе достаточно высока. Потенциостат представляет собой электронную схему, используемую в электрохимии для изучения электрохимических процессов, происходящих на определённом электроде и одна из сложностей – электрохимические исследования часто требуют настройки условий для проведения конкретных исследований. Для обнаружения следов глюкозы были использованы контакты лазер-индуцированной графеновой пены.

Простая схема потенциостата использует усилитель управления СА, который сравнивает измеренное напряжение ячейки с желаемым и подаёт ток для достижения одинакового напряжения в инвертирующей конфигурации с отрицательной обратной связью [1]. Повторитель напряжения VF измеряет напряжение RE и передаёт сигнал в контур обратной связи. Трансимпедансный усилитель TIA преобразует измеренный ток в WE и напряжение через сопротивление R_f . Современные усилители VF имеют близкие к нулю входные токи, что позволяет игнорировать этот эффект [2]. Заземление WE обеспечивает измерение напряжений относительно земли (Рис1).

Платы микроконтроллеров Arduino на основе семейства микроконтроллеров ATmega выделяются своими выдающимися возможностями, доступной ценой и портативностью. Кроме того, эти платы имеют цифровые выходы до 5 В с диапазоном $\pm 2,5$ В, что позволяет этим платам создавать требуемый диапазон выходного напряжения для электрохимических экспериментов с помощью схемы сдвига напряжения [3].

Для выполнения циклической вольтамперометрии к противоэлектроду должна быть применена треугольная кривая выходного напряжения. Для выполнения такого сканирования выходного напряжения была реализована ступенчатая функция с использованием двух циклов for, один из которых увеличивает (прямое направление цикла), а другой уменьшает (обратное направление цикла) значение в функции analogwrite() в каждом цикле, как показано на рисунке 5. Функция analogwrite() будет увеличивать и уменьшать, соответственно, рабочий цикл выходного контакта ШИМ. Затем сигнал

фильтруется, и смещение добавляется схемой потенциостата перед его применением к противоэлектроду.

В рамках работы по разработки прибора для неинвазивного измерения уровня глюкозы в крови была произведена разработка и тестирование простого, экономически эффективного потенциостата на базе Arduino. Было продемонстрировано, что он может выполнять электрохимические измерения циклической вольтамперометрии и потенциальной ступенчатой вольтамперометрии. Также с помощью программного обеспечения можно настраивать скорость сканирования измерений циклической вольтамперометрии. Полученные данные через последовательный интерфейс USB можно отображать в реальном времени с использованием среды программного обеспечения обработки с открытым исходным кодом в качестве визуального интерфейса последовательного терминала, который также может сохранять данные в текстовом формате файла для анализа.

Было произведено три эксперимента с циклическим измерением вольтамперметрии, при каждом при каждом в раствор NaOH был добавлен 1 грамм сахара. Результаты измерений отображены на общем графике (Рис3). Данный программно-аппаратный комплекс показал высокую чувствительность даже к малым изменениям уровня глюкозы в растворе. Следующим этапом будет проведение эксперимента с большим количеством измерений и построение кривой зависимости результатов измерения от концентрации глюкозы и анализ точности измерений методов наложения их на сетки кларка [4].

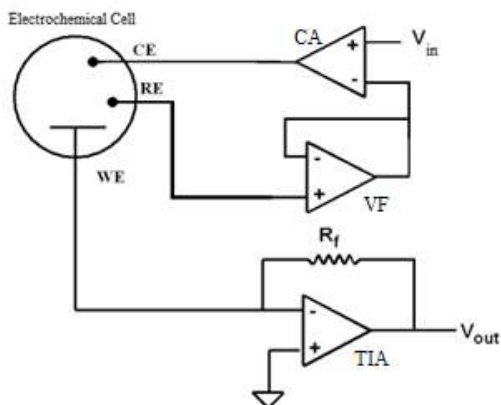


Рис.1 – Простая схема потенциостата


```

//Cyclic Voltammetry function
void loop() {
    int maxvoltage = 120;//voltage range
    int voltage_scan = 1;
    delay(2000);//waits to stabilize the voltage values
    for (int i=0; i <maxvoltage; i=i+voltage_scan){
        analogWrite(3, i);
        findAverage();
    }
    for (int i=maxvoltage; i>0; i=i-voltage_scan){
        analogWrite(3,i);
        findAverage();
    }
    delay(2000);//waits to stabilize the voltage values
    analogWrite(2,LOW);
    Serial.flush();
    while(1);//leaves arduino in a invisible loop
}

```

Рис.2 – Программный код циклической обработки

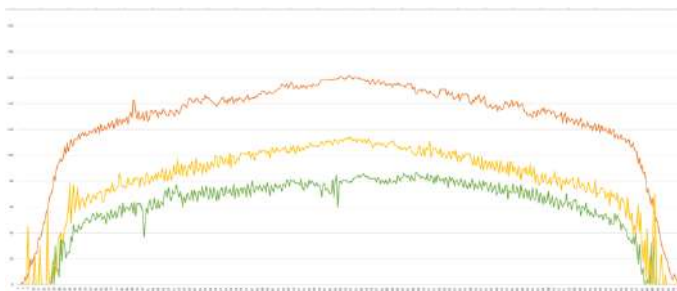


Рис.3 – Результаты циклической вольтампертметрии трех измерений при повышении концентраций глюкозы

Литература:

1. Dryden M.D., Wheeler A.R. DStat: A Versatile, Open-Source Potentiostat for Electroanalysis and Integration // PLoS One. – 2015. – Oct. 28;10(10). – doi: 10.1371/journal.pone.0140349.
2. Meloni G.N. Building a microcontroller based potentiostat: An inexpensive and versatile platform for teaching electrochemistry and instrumentation // J. Chem. Educ., 2016. – Vol. 93, N. 7. – P. 1320–1322. – doi: 10.1021/acs.jchemed.5b00961.

3. Kaswan K.S., Singh S.P., Sagar S. Role of Arduino in real world applications // Int. J. Sci. Technol. Res. – 2020. – Vol. 9, N. 1. – P. 1113–1116.

4. Kovatchev B.P., Wakeman C.A., Breton M.D., Kost G.J., Louie R.F., Tran N.K., Klonoff D.C. Computing the surveillance error grid analysis: Procedure and examples // J. Diabetes Sci. Technol. – 2014. – Vol. 8. – P. 673–684. – doi: 10.1177/1932296814539590.

ДИСКРЕТНАЯ СИСТЕМА СТАБИЛИЗАЦИИ СКОРОСТИ ВРАЩЕНИЯ ВАЛА ДВИГАТЕЛЯ ПОСТОЯННОГО ТОКА С НЕЗАВИСИМЫМ ВОЗБУЖДЕНИЕМ

Р.П. Парыгин

**Новосибирский государственный технический университет,
г. Новосибирск, parygin154@gmail.com**

Научный руководитель: Саблина Г.В., к.т.н., доцент

В данной статье рассмотрены три наиболее распространённых метода получения дискретных моделей непрерывных систем: матричный метод, z-преобразование и преобразование Тастина. Исследование методов произведено на примере объекта «двигатель постоянного тока с независимым возбуждением». Разработаны регуляторы для стабилизации скорости вращения двигателя на основе модального метода. Представлены результаты моделирования системы, полученные в программной среде Matlab Simulink.

This article discusses the three most common methods for fragment models of continuous systems: the matrix method, the z-transform and the Tustin transform. The research methodology was carried out using the example of a “DC motor with independent connection” object. Regulators have been developed to ensure the performance of an intelligent engine based on the modal method. The results of a system study obtained in the Matlab Simulink software environment are presented.

Практически все физические процессы, окружающие человека, происходят на непрерывной временной шкале. Для обработки таких данных используют вычислительную технику, работа с которой подразумевает преобразование сигналов в дискретную форму. Это необходимо для функционирования таких систем, как радио, навигационные системы, системы оптической связи, системы управления двигателями.

В данной работе получены дискретные модели объекта управления «двигатель постоянного тока с независимым возбуждением» тремя способами, и реализованы регуляторы на основе модального метода

синтеза. Схема управления двигателем постоянного тока с независимым возбуждением представлена на рис.1.

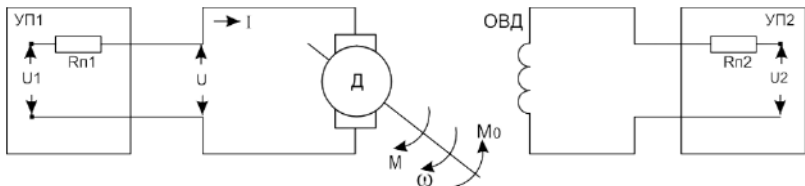


Рис.1 – Схема управления двигателем постоянного тока с независимым возбуждением

где: ОВД – обмотка возбуждения двигателя; Д – двигатель постоянного тока; УП1, УП2 – управляемые преобразователи для питания якорной цепи и обмотки возбуждения двигателя; U_1 , U_2 – напряжения на входах управляемых преобразователей; $R_{П1}$, $R_{П2}$ – активные сопротивления преобразователей; U , I – напряжение и ток на якоре двигателя; ω – угловая скорость вращения вала двигателя; M – момент, развиваемый двигателем; M_c – момент сопротивления на валу двигателя, который будем считать внешним возмущением.

Математическая модель объекта в линейном приближении [3] имеет вид:

$$\begin{cases} J \frac{d\omega}{dt} = C\Phi_H I - M_c, \\ L \frac{dI}{dt} = -RI - C\Phi_H \omega + U_1. \end{cases} \quad (1)$$

Коэффициенты системы рассчитаны на основе паспортных данных, которые приведены в таблице 1.

Таблица 1 – Параметры двигателя

Параметры двигателя										
P_H , [кВт]	U_H , [В]	I_H , [А]	ω_H , [рад/с]	η	$P_{П}$	β	J_d , [кг/м ²]	$t_{П}$, [с]	σ , %	$\frac{\omega_m}{\omega_H}$, %
0,46	220	4,8	110	0,63	2	0,25	0,018	0,46	20	4,8

Передаточная функция объекта определяется с учётом передаточной функции тахогенератора в виде:

$$W_{ov}(p) = W_o(p) \cdot W_{tx}(p) = \frac{62,5606}{(p^2 + 326,421p + 2426,656)} = \frac{B(p)}{A(p)} \quad (2)$$

Таблица 2. – Параметры тахогенератора

Далее рассматриваются способы перехода от непрерывной передаточной функции к её дискретной модели.

Матричный метод

Переход к дискретной модели матричным методом осуществляется за счет процедуры разложения в ряд Тейлора (общее решение системы линейных дифференциальных уравнений).

Дискретная модель объекта, полученная матричным методом:

$$W(z) = \frac{0,0013311508z + 0,000480412}{z^2 - 0,96795436z + 0,03822288} \quad (3)$$

Для оценки качества переходных процессов, в программной среде Matlab Simulink при шаге дискретизации выполнено моделирование непрерывного и дискретного объектов. Шаг дискретизации: $T = 0,01$ с.

Преобразование Тастина

Для дифференциального оператора p выполним преобразование:

$$p = \frac{2}{T} \cdot \frac{z-1}{z+1}, \text{ где } z - \text{оператор сдвига.}$$

Наименование тахогенератора	Параметры тахогенератора				
	E_y	$n_{max},$ [об/мин]	P_y	I_{nmax}	R_y
ТД-103	6,6	1500	0,041	0,1	660

Передаточная функция в дискретном виде:

$$W(z) = \frac{0.00058081z^2 + 0.00116163z + 0.00058081}{z^2 - 0.69767051z - 0.21221207} \quad (4)$$

Метод z-преобразования

Дискретная модель передаточной функции имеет вид:

$$W(z) = \frac{z-1}{z} Z \left\{ \frac{W_o(p)}{p} \right\}$$

Передаточная функция, полученная методом z-преобразования:

$$W(z) = \frac{0.0013305970603z + 0.000480402}{z^2 - 0.967958941952 + 0.0382271242101} \quad (5)$$

Произведём сравнение переходных процессов непрерывного и дискретных объектов (рис. 2).

Графики переходных процессов для матричного метода и метода z-преобразования совпадают. При малом шаге дискретизации, а в нашем случае $T = 0,01$ с, эти методы показывают очень высокую точность.

Если обратить внимание на результат работы метода преобразования Тастина, можно заметить, что график переходного процесса не совпадает с непрерывной моделью. В начальный момент времени переходный процесс в дискретной модели имеет значение, отличное от нуля $y(0)=0,0006$. Так получается из-за того, что числитель и знаменатель дискретной модели, имеют одинаковый порядок, чего не наблюдается в двух вышеописанных методах.

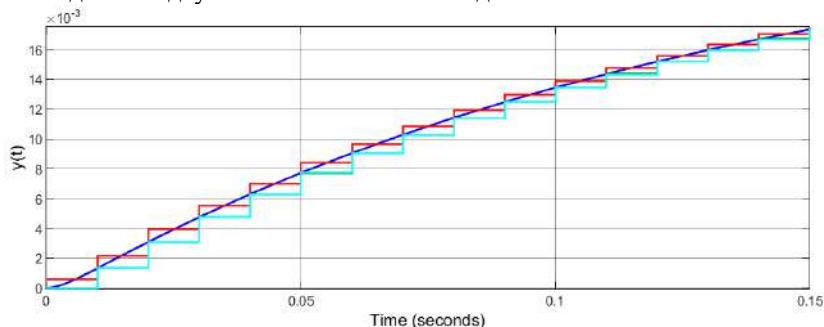


Рис. 2 – Графики переходных процессов непрерывного объекта (синий) и дискретных объектов, полученных методами: матричный (зелёный), Z-преобразование (голубой) и Тастина (красный).

Реализация дискретных регуляторов

Для возможности влиять на качество переходных процессов, а также реакции системы на внешнее возмущение, для всех полученных дискретных моделей были реализованы дискретные регуляторы с использованием модального метода синтеза.

Для оценки работоспособности произведено моделирование.

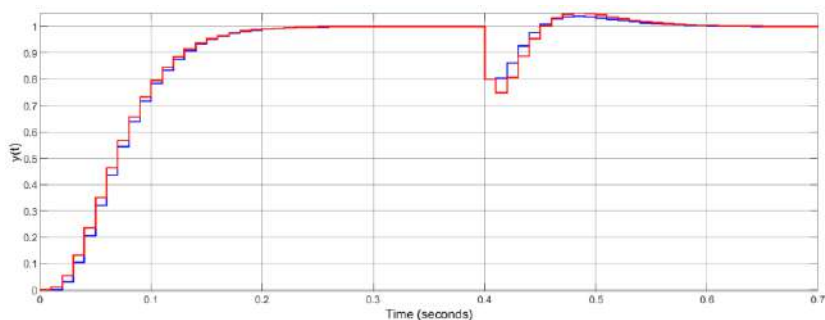


Рис. 4 – Переходные процессы с возмущением для синтезированных систем на основе дискретных объектов, полученных Z-преобразованием (синий график), матричным методом (зеленый график) и преобразованием Тастина (красный график)

Возмущение, приложенное к выходу дискретного объекта, отрабатывается во всех случаях. Модальный метод синтеза дискретного регулятора применим для дискретных объектов, полученных любыми методами.

Графики переходных процессов совпадают для систем, полученных Z-преобразованием и матричным методом.

Дискретный регулятор, рассчитанный на основе дискретной модели, полученной преобразованием Тастина, показал непригодность в работе с непрерывным объектом.

Литература:

1. Волков Н.И., Миловзоров В.П. Электромашинные устройства автоматики: учебник для вузов. – М.: Высшая школа, 1986. – 336 с.
2. Саблина Г.В. Цифровые системы управления. Сборник задач для индивидуальных заданий: учеб. пособие, Изд-во НГТУ, 2019. – 66 с.
3. Французова Г.А. Расчет и исследование системы регулирования динамическим объектом: учебно-методическое пособие, Изд-во НГТУ, 2019. – 44 с.

РАЗРАБОТКА ИНСТРУМЕНТАЛЬНОГО КОМПЛЕКСА ДЛЯ ПОСТРОЕНИЯ И ИССЛЕДОВАНИЯ АНАЛОГОВЫХ И ЦИФРОВЫХ УСТРОЙСТВ С ВОЗМОЖНОСТЬЮ ОБРАБОТКИ СИГНАЛОВ ИЗМЕРИТЕЛЬНЫХ ПРИБОРОВ

Т.А. Абдукаримов, А.В. Ковальчук

**Новосибирский государственный технический университет,
г. Новосибирск, abdukarimovtolib12@gmail.com**

**Научные руководители: Овчеренко В.А., к.т.н., доцент,
Лаптев Д.В., к.т.н.**

В статье рассматривается необходимость разработки инструментального комплекса для проектирования и анализа аналоговых и цифровых устройств с возможностью обработки сигналов измерительных приборов. Обоснована актуальность создания отечественного программного обеспечения для образовательных учреждений в условиях импортозамещения, проведен анализ существующих аналогов и обозначены их недостатки. Предложено новое решение, которое будет доступным и интуитивно понятным для пользователей.

The article addresses the need to develop a toolset for the design and analysis of analog and digital devices with the capability to process signals from measuring instruments. The importance of creating domestic software for educational institutions in the context of import substitution is substantiated, and an analysis of existing analogs is provided, highlighting their limitations. A new solution is proposed that will be accessible and intuitive for users.

В условиях импортозамещения остро стоит необходимость создания отечественного программного комплекса для построения и исследования аналоговых и цифровых устройств с возможностью обработки сигналов измерительных приборов, особенно для учебных заведений. Существующие зарубежные решения, такие как MATLAB, LabVIEW, Multisim и Proteus, обладают высокой функциональностью, но из-за их сложности, высокой стоимости лицензий и требований к оборудованию их использование ограничено [1, 2, 3, 4]. Это является значительной проблемой, что подчеркивает необходимость создания нового решения, соответствующего потребностям образовательного процесса.

Предлагаемый инструментальный комплекс представляет собой доступное и интуитивно понятное программное обеспечение, ориентированное на моделирование и исследование аналоговых и цифровых устройств. Комплекс включает основные компоненты для построения электронных схем, что позволяет пользователям моделировать разнообразные схемы и процессы, необходимые для

учебных и исследовательских целей. На рисунке 1 представлен пользовательский интерфейс инструментального комплекса.

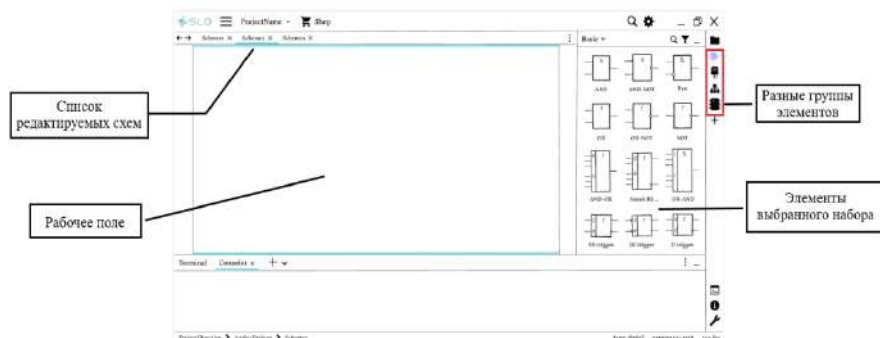


Рис. 1 – Пользовательский интерфейс инструментального комплекса

Процесс моделирования будет происходить в формате визуального программирования. Пользователи смогут использовать графический интерфейс для построения схем, перетаскивая существующие и созданные из более простых компонентов элементы и соединяя их с помощью виртуальных проводов. Такой подход упрощает создание моделей, позволяя сосредоточиться на логике и структуре схемы, что особенно полезно для начинающих пользователей и образовательного процесса. Предусмотрена обработка двух типов сигналов: цифровых и аналоговых. В связи с этим все элементы делятся на три условные группы: цифровые, аналоговые и цифро-аналоговые. Цифровые элементы предназначены для работы исключительно с цифровыми виртуальными сигналами, аналоговые — только с аналоговыми, а цифро-аналоговые обеспечивают взаимодействие с обоими типами сигналов, что позволяет интегрировать их в более сложные модели с элементами обоих типов.

Предлагаемые наборы цифровых компонентов охватывают логические вентили (И, ИЛИ, НЕ, И-НЕ, ИЛИ-НЕ, ИЛИ-исключающее и другие), запоминающие и сдвиговые регистры, мультиплексоры, демультиплексоры, шифраторы, дешифраторы, триггеры (RS, D, T), счётчики различной разрядности (2, 4, 8, 16 и 32), а также источники цифрового сигнала и элементы визуализации, такие как 7-сегментные дисплеи, пиксели и пиксельные матрицы. Эти компоненты обеспечивают полноценную поддержку работы с цифровыми сигналами.

Для работы с аналоговыми сигналами в систему будут включены резисторы и конденсаторы с настраиваемыми номиналами,

операционные усилители, усилитель с неравномерной АЧХ и др., осциллограф, транзистор, диод, катушки индуктивности, мультиметр, лампа и источники сигналов. Эти устройства позволяют моделировать основные элементы аналоговых цепей, обеспечивая широкие возможности для исследования аналоговой электроники. Кроме того, в комплекс будут добавлены элементы, способные моделировать работу как с аналоговыми, так и с цифровыми сигналами, такие как АЦП и ЦАП.

Предлагается возможность считывания данных с платы Arduino UNO, подключенной к внешнему порту вычислительного устройства. Поддержка считывания данных с внешних устройств позволит интегрировать измерительные приборы с инструментальным комплексом для цифровой обработки их сигналов. В будущем планируется расширение списка поддерживаемых внешних устройств. Изначально система будет поддерживать Arduino UNO, но со временем список будет дополнен. Это обеспечит комплексную работу с внешними измерительными приборами (Рисунок 2).

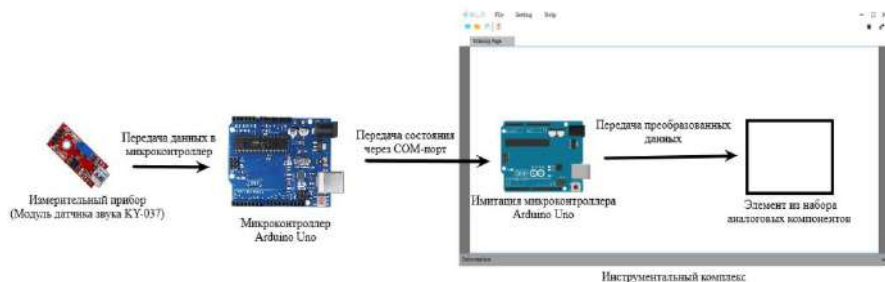


Рис. 2 – Схема передачи сигнала от измерительного прибора в инструментальный комплекс

Предлагаемый инструментальный комплекс будет разработан с возможностью масштабирования и расширения функционала. В будущем это позволит добавлять в него новые цифровые и аналоговые компоненты, повышая гибкость и расширяя возможности системы.

Литература:

1. Техническая документация LabVIEW [Электронный ресурс] // URL: <https://www.ni.com/docs/en-US/> (дата обращения: 10.09.2024).
2. Техническая документация NI Multisim [Электронный ресурс] // URL: <https://www.ni.com/docs/en-US/bundle/multisim/resource/374483d.pdf> (дата обращения: 27.09.2024).

3. Proteus Helper page [Электронный ресурс] // URL: <https://www.labcenter.com/downloads/#help> (дата обращения: 02.10.2024).

4. Техническая документация Matlab [Электронный ресурс] // URL: <https://www.mathworks.com/help/matlab/> (дата обращения: 10.10.2024).

ЛАБОРАТОРНАЯ АКУСТИЧЕСКАЯ КАБИНА НА ОСНОВЕ GB1.5 И ЗВУКОИЗОЛЯЦИИ АКУСТИКНАУФ

В.М. Куршин, М.А. Чириков
Новосибирский государственный технический университет,
г. Новосибирск, max_chir@mail.ru

Работа посвящена созданию акустической кабины для лаборатории с использованием вибродемпфирующих, звукоизолирующих и акустических материалов, и протестированную с помощью программно-реализованного шумомера.

The work is devoted to the creation of an acoustic cabin for a laboratory using vibration damping, soundproofing and acoustic materials, and tested using a software-implemented noise meter.

С развитием технологий и общей урбанизацией человечество столкнулось с очередным фактором, способным оказать негативное влияние на здоровье организма — техногенным шумом. Вредное воздействие шума может выражаться через приобретение стресс-зависимых заболеваний, повреждению органа слуха, снижению когнитивных функций и т. д. [1].

Для противодействия шумовому загрязнению окружающей среды немаловажную роль играет подготовка специалистов, способных выполнять необходимые измерения и разрабатывать решения для поставленных задач. В рамках данной работы для образовательных целей была создана лабораторная акустическая кабина, предоставляющая все необходимые условия для проведения измерений в рамках образовательного курса.

При создании акустической кабины использовались специальные материалы, перечисленные в таблице 1.

Помимо специальных материалов в процессе сборки акустической кабины на кафедре Защиты информации НГТУ использовались листы фанеры, деревянные бруски и различный крепеж.

Стоимость готовых решений на рынке варьируется от 80 тыс. до нескольких миллионов, что зависит от размеров и используемых

материалов. В рамках данной работы стояла задача создать рабочую акустическую кабину с минимально возможными затратами. Часть материалов была получена в столярной мастерской, работы велись сотрудниками кафедры, а доставка материалов осуществлялась на личном автотранспорте сотрудников. Общие затраты составили около 40 тыс. рублей.

Таблица 1 – Материалы акустической кабины

Наименование, размеры	Общие характеристики
Вибродемпфирующий материал GB1.5, 350x570x15 мм	Основной материал: полимер Общий индекс звукоизоляции — 25 дБ
Звукоизоляция АкустиКнауф, 610x100x50 мм	Основной материал: стекловолокно Эффективное снижение шума: 52 дБ
Акустический поролон «Пирамида», 1000x2000 мм	Основной материал: пенополиуретан Толщина: 40 мм пирамида, 20 мм основание

Технология постройки лабораторной акустической кабины:

1. Фанерные заготовки (190x200x0.3 см, 12 шт.) и деревянные бруски (5x5 см) скрепляются для получения стенки куба;
2. Плита «АкустиКнауф» нарезается канцелярским ножом и вкладывается в ниши стенок, еще одна фанера устанавливается сверху, образуя сэндвич-панель;
3. Из 6 стенок собирается куб, одна из которых крепится на петли;
4. Внутренняя поверхность заклеивается вибродемпфирующим материалом «GB1.5» с помощью прикаточного ролика;
5. Поверх вибродемпфирующего материала при помощи специального клея и строительного степлера наклеивается акустический поролон;
6. С одной стороны кабины продельвается техническое отверстие для проведения кабелей и проводов.

Габаритные размеры кабины указаны в таблице 2. Собранная акустическая кабина представлена на рис. 1.

Таблица 2 – Габаритные размеры акустической кабины

Внешние размеры	200x190x190 см
Толщина стенок	11,5 см
Внутренние размеры	188,5x178,5x178,5 см



Рис. 1 – Акустическая кабина

Для тестирования звукопоглощающих свойств акустической кабины были проведены тестовые испытания с использованием измерительной системы, состав которой приведен в таблице 3.

Таблица 3 – Состав измерительной системы

Наименование	Характеристики
Микрофон ZT-333 с предуслителем P410	Номинальная чувствительность: 48,14 мВ/Па Частотный диапазон: 25 — 20000 Гц
АЦП-модуль LTR-24	Частота преобразования: 610 Гц — 117 кГц Разрядность: 20/24 бит
ПЭВМ с предустановленным ПО	Ноутбук с ОС Windows 10, специализированное ПО и библиотеки от производителя L-CARD

Общая схема измерительной системы составлялась согласно перечню рекомендаций и требований [2 — 4] и приведена на рис. 2.

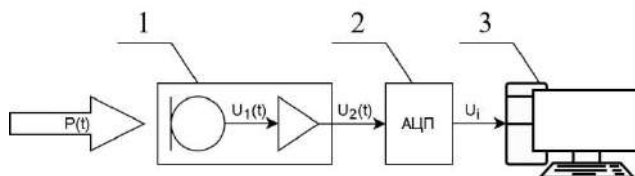


Рис. 2 – Схема оборудования (1 — микрофон с предуслителем; 2 — АЦП-модуль; 3 — ПЭВМ с ПО)

Также в ходе испытаний использовалась акустическая колонка, через которую генерировался тестовый сигнал «белый шум» (БШ). Запись и обработка сигнала проводилась через измерительную систему (рис. 2), контроль над системой выполнялся через разработанную программу на платформе LabVIEW.

Была поставлена задача определить общий интегральный уровень тестового сигнала и его интегральные уровни в отдельных частотных полосах при различных условиях. Пример полученных результатов измерения с интерфейсом программы представлен на рис. 3.

В ходе экспериментов были получены общие интегральные уровни измеренного звукового сигнала при различных условиях. Результаты измерения представлены в таблице 4. Акустическая колонка располагалась внутри кабины, а микрофон был установлен снаружи.

Таблица 4 – Результаты измерения

Условия эксперимента	Общ. интегральный уровень L, дБ
1. Закрытая кабина, без БШ	50,4257
2. Открытая кабина, без БШ	53,5871
3. Закрытая кабина, БШ 60 дБ	52,1321
4. Открытая кабина, БШ 60 дБ	60,3467

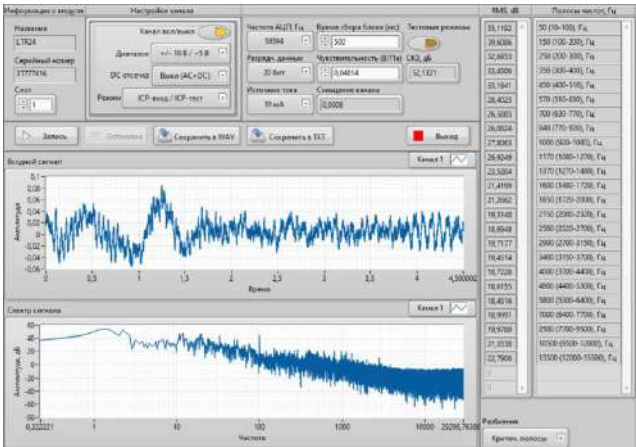


Рис. 3 – Интерфейс программы с результатами измерения тестового сигнала (камера — закрыта; колонка — внутри; микрофон — снаружи)

Из таблицы 4 можно наблюдать, что разница между измеренными значениями уровня L при БШ 60 дБ для закрытого и открытого состояний кабины составляет более 8 дБ. Когда источником сигнала

являются фоновые шумы (работа системы освещения, отопления, кулеров ПЭВМ, различные звуки за пределами лаборатории и т. п.), то разница составляет около 3 дБ. На рис. 4 продемонстрированы графики зависимости L от частотной полосы для 4-х вышеуказанных экспериментов.

Можно наблюдать, что для первых 3-х экспериментов разница между L на практически всех частотных полосах незначительна и составляет примерно 1 – 1,5 дБ. Эксперимент 4 проводился с открытой камерой и БШ 60 дБ, по соответствующему графику можно наблюдать значительную разницу в L относительно первых 3-х экспериментов, на частоте 13,5 кГц разница составляет около 29,5 дБ, что наглядно демонстрирует звукопоглощающие свойства акустической кабины.

В заключении также можно отметить, что, начиная от низких частот и до 1,6 кГц (рис. 3) эффективность звукоизоляции несколько ниже, причем спад происходит в сторону низких частот. Данный эффект можно объяснить особенностью пирамидального акустического поролона, который считается наиболее эффективным для средних и высоких частот [5].

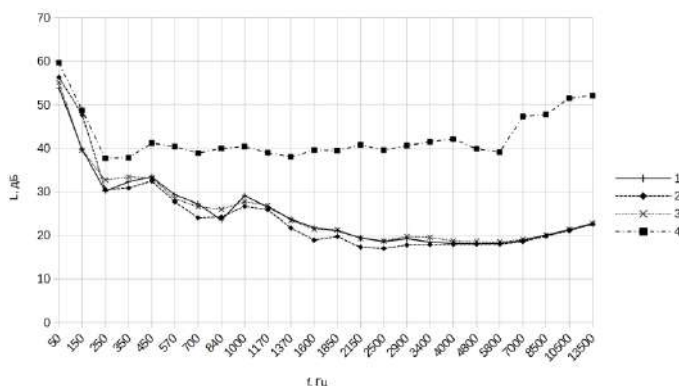


Рис. 4 – Графики зависимости L от частотной полосы (графики 1, 2, 3, 4 соответствуют № эксперимента из таблицы 3)

Литература:

1. Романов С.Н. Биологическое действие вибрации и звука: Парадоксы и проблемы XX века. – Л.: Наука, 1991. - 158 с.
2. Иванов А.В. Оценка защищенности информации от утечки по виброакустическим каналам: учебное пособие / А. В. Иванов. - Новосибирск: Изд-во НГТУ, 2018. - 76 с.

3. Трушин В.А. Защита речевой информации от утечки по акустическим и виброакустическим каналам: учеб. пособие / В.А. Трушин. – Новосибирск, 2006. 40 с.

4. Хорев А.А. Оценка эффективности систем виброакустической маскировки / Ю. К. Макаров, А.А. Хорев. – 18 с.

5. Виды акустического поролонa [Электронный ресурс] / ЛюксПолимеру – URL: <https://www.luxpolymer.ru/stati/vidyi-akusticheskogo-porolona/> (дата обращения 06.11.2024)

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ЧАСТОТНО-ЗАВИСИМОГО УСИЛИТЕЛЯ ШУМОМЕРА В СРЕДЕ РАЗРАБОТКИ PYTHON

В.Д. Хозяева, М.А. Чириков

**Новосибирский государственный технический университет,
г. Новосибирск, valerochka2001@mail.ru**

Научный руководитель: Трушин В.А., к.т.н., с.н.с.

Настоящая работа посвящена созданию модели частотно-зависимого усилителя шумомера на основе его реализации в среде разработки Python

This work is devoted to the development of the model of the frequency dependent amplifier of sound level meter on its software implementation in the Python development environment.

В данной статье рассматривается разработка модели частотно-зависимого усилителя шумомера с использованием языка программирования (ЯП) Python, используемые библиотеки, функционал разработанного программного обеспечения (ПО) и планируемые «расширения» возможностей ПО [1].

Выбор данного ЯП обосновывается следующими преимуществами: бесплатность, широкий выбор библиотек для работы с данными, кроссплатформенность, возможность совмещать ПО с другими языками программирования.

Так как шумомер является упрощенной реализацией модели уха, то параметры нашего ПО подбираются относительно параметров слухового аппарата человека. Первым этапом обработки звукового сигнала является его частотно-зависимое усиление [2]. Значительное усиление происходит на частотах в промежутке от 2 КГц до 5 КГц [3].

ЯП Python обладает возможностью, позволяющей с помощью одной функции выполнить преобразование Фурье [4]. Для этого существует библиотека `scipy` (библиотека для сложных научных

вычислений и анализа данных) с модулем fft (модуль, в котором находятся функции для выполнения дискретного преобразования Фурье) [5]. После разложения входного сигнала на гармоники, необходимо их усилить. При выборе функции усиления необходимо учитывать амплитудно-частотную характеристику (АЧХ) слухового канала человека. Данная АЧХ с учётом совместного действия ушной раковины и слухового канала показана на Рис. 1 [6].

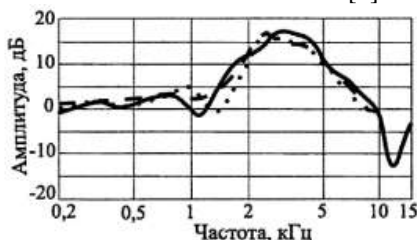


Рис. 1 - АЧХ слухового канала у барабанной перепонки [6]

За основу для построения АЧХ цифрового частотно-зависимого усилителя использована скорректированная функция нормального распределения для программной реализации «колоколообразной» функции в частотной области усиления АЧХ, что позволит с вручную подобранными параметрами смоделировать усиление средних частот слухового канала. Функция АЧХ определяется формулой (1).

$$AЧХ = \frac{5 \cdot 10^3}{D \sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2D^2}} + 1, D = 3500, \mu = 600 \quad (1)$$

АЧХ разработанного частотно-зависимого усилителя показана на Рис.2. Усиление в дБ представлено графиком на Рис.3

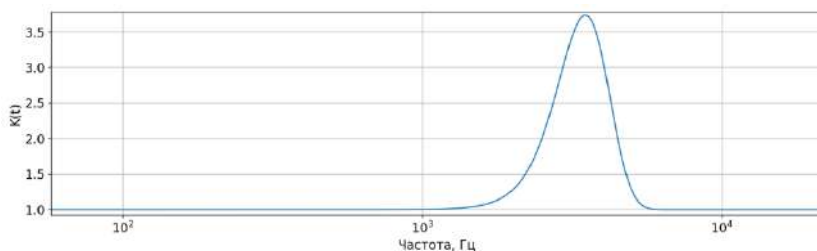


Рис. 2 - АЧХ частотно-зависимого усилителя

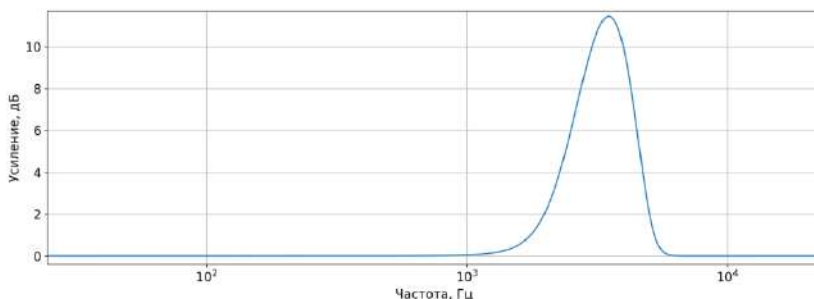


Рис. 3 - АЧХ усилителя в децибелах

Заметим, что в результате разработанный усилитель имеет наибольшее усиление в частотном интервале 2 – 5 кГц, что и необходимо согласно анатомии слухового канала наружного человеческого уха.

Блок-схема разработанной модели цифрового усилителя представлена на Рис 4.

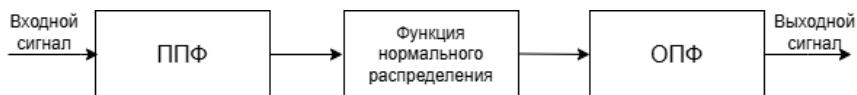


Рис. 4 - Блок-схема модели цифрового усилителя

Удостоверимся в том, что разработанный усилитель изменяет входной сигнал согласно полученной для него АЧХ. Для этого сгенерируем два тональных сигнала: 750Гц и 3000Гц. Исходные сигналы и усиленные сигналы представлены на Рис. 5 и Рис. 6.

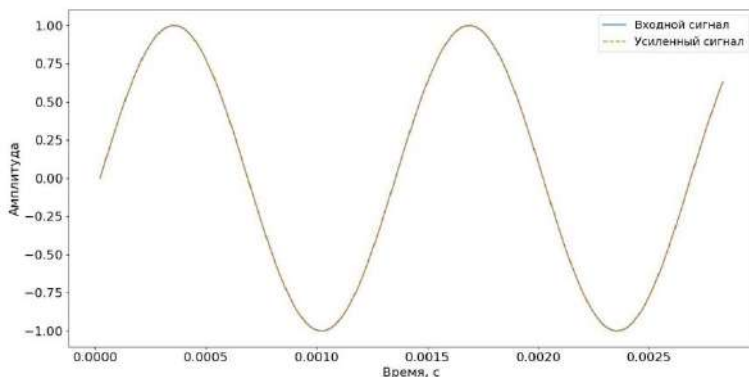


Рис. 5 - Входной и выходной сигналы на частоте 750Гц

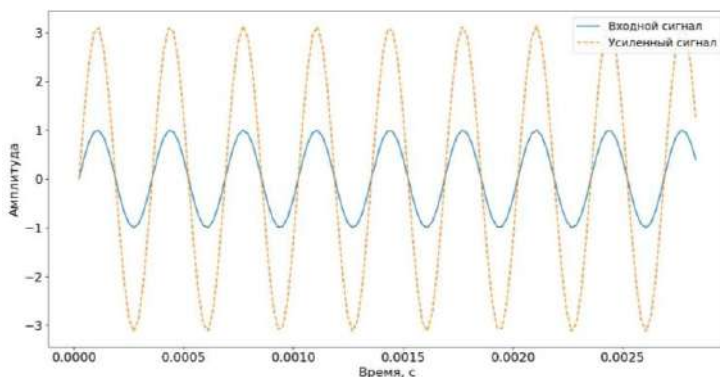


Рис. 6 - Входной и выходной сигналы на частоте 3КГц

Дальнейшей целью разработки является проектирование фильтра и преобразователя сигнала усилителя, а также создание ПО для вывода результатов.

Литература:

1. Аппаратно-программная реализация ИИ модели слуховой системы человека в среде LabVIEW / Д. В. Бухтоярова, С. А. Рожков, М. А. Чириков, В. А. Трушин. - Текст: электронный // Наука. Технологии. Инновации: сб. науч. тр. 15 Всерос. науч. конф. молодых ученых, посвящ. Году науки и технологий в России, Новосибирск, 6–10 дек. 2021 г.: в 10 ч. – Новосибирск: Изд-во НГТУ, 2021. – Ч. 1. – С. 52–55.
2. ГОСТ Р 53188.1 – 2019. Шумомеры. Часть 1. Технические требования. – Введ. 2019-04-23. – М.: Стандартинформ, – 2019. – 46 с.

3. Алдошина И., Основы психоакустики [Электронный ресурс] / Библиотека-гитариста.рф. – URL: <https://xn----7sbbabaaqnzbebul0c0bsige.xn--p1ai/knigi/osnovy-psihoakustiki-aldoshina-i/> (дата обращения: 06.11.2024).
4. Лукин А., Введение в цифровую обработку сигналов (математические основы). – Лаборатория компьютерной графики и мультимедиа, МГУ, 2007. – 54 с.
5. SciPy [Электронный ресурс] // SciPy: [сайт]. – URL: <https://scipy.org/> (дата обращения: 06.11.2024).
6. Алдошина, И. А. Музыкальная акустика: Учебник / И. А. Алдошина, Р. Приттс. – Санкт-Петербург: ООО Издательство «Композитор», 2006. – С. 39-196.

АНАЛИЗ УЯЗВИМОСТЕЙ В ПРОГРАММНОМ КОДЕ С ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Б.Л. Патрушев

**Новосибирский национальный исследовательский
государственный университет,**

г. Новосибирск, b.patrushev@ngs.ru

Научный руководитель: Пестунова Т.М., к. т. н., доц.

Объектом исследования является процесс анализа уязвимостей в программном коде на языке Python с использованием методов машинного обучения (МО). Задачами работы являются: исследовать анализ уязвимостей в коде на языке Python с использованием МО, а также собрать реальный набор данных с уязвимостями на языке Python.

The object of the study is the process of vulnerability analysis in Python program code using machine learning (ML) methods. The objectives of the work are: to study the analysis of vulnerabilities in Python code using ML, as well as to collect a real data set with vulnerabilities in the Python language.

В современном мире безопасность программного обеспечения является критически важной. Уязвимости в коде могут привести к серьезным последствиям, включая утечку данных и нарушение функционирования систем. Применение методов машинного обучения для анализа уязвимостей представляет собой новый подход к автоматизации обнаружения потенциальных угроз.

Подход, основанный на машинном обучении, представляет собой области анализа уязвимостей, позволяющий обнаружить проблемы, которые могли бы остаться незамеченными при использовании

традиционных методов. В ходе анализа существующих исследований, не было найдено подходящих решений поставленной задачи [1].

Цель работы – разработка и анализ эффективности методов обнаружения уязвимостей в коде Python, основанных на машинном обучении, создание набора данных из уязвимого программного кода, собранных с открытых репозиторийев.

В ходе работы были проанализированы существующие решения, разработаны: генератор ключевых слов; скрипт на языке Python позволяющий по конкретному номеру CVE получать номера CVE; инструмент для сбора набора данных по ключевым словосочетаниям в сообщении коммита в репозиториях GitHub на языке Python с использованием распределенной очереди задач Celery и фильтрацией статическим анализатором кода CodeQL и получен набор данных с помощью этого инструмента; пайплайн для нахождения двух типов уязвимостей - SQL-инъекция и инъекцией команд с использованием архитектуры LSTM, где был получен результат F-меры в 84.6%. По итогам исследования подтверждена возможность использования машинного обучения для нахождения уязвимостей в коде и получены оценки эффективности предложенного подхода. Результат работы может быть использован в других смежных исследованиях.

Литература:

1. *Marjanov T. Machine Learning for Source Code Vulnerability Detection: What Works and What Isn't There Yet // IEEE Security and Privacy, –2022. – Sep. – Vol. 20, N. 05. – P. 60–76.*

ВОССТАНОВЛЕНИЕ ПЕРИОДИЧЕСКИХ СИГНАЛОВ ПРИ АНАЛОГОВО-ЦИФРОВОМ ПРЕОБРАЗОВАНИИ

В.А. Серебренников

Новосибирский государственный технический университет,

г. Новосибирск, s3rebrebrennikov38@yandex.ru

Научный руководитель: Д.В. Лаптев, к.т.н.

Настоящая работа посвящена исследованию метода восстановления периодических сигналов при аналогово-цифровом преобразовании

This work is devoted to the study of the method of restoring periodic signals during analog-to-digital conversion.

Необходимость создания метода восстановления периодических сигналов при аналогово-цифровом преобразовании, основанного на

свойстве периодических сигналов повторятся во времени, обусловлена важностью корректного решения многих практических задач в электронике, средства связи, технической защите информации и др. [1,2].

В данной работе рассматривается модель, математический аппарат которой схож с аппаратом, используемый в китайской теореме об остатках. Для восстановления формы сигнала из отсчёта, взятого в случайный момент времени, вычитается целое число периодов и используется «остаток» периода, другими словами, временная координата точки делится по модулю периода восстанавливаемого. Временная координата точки определяется формулой (1).

$$\begin{aligned} \bar{t}_1 &= t_1 \bmod T \\ \bar{t}_2 &= t_2 \bmod T \\ \bar{t}_3 &= t_3 \bmod T \\ &\dots \\ \bar{t}_n &= t_n \bmod T \end{aligned} \quad (1)$$

Далее к полученным данным применяется аппроксимация, тем самым получая восстановленный сигнал.

В процессе измерения необходимо получить период (частоту) измеряемого сигнала, и набор точек с двумя координатами: амплитуда, временная метка.

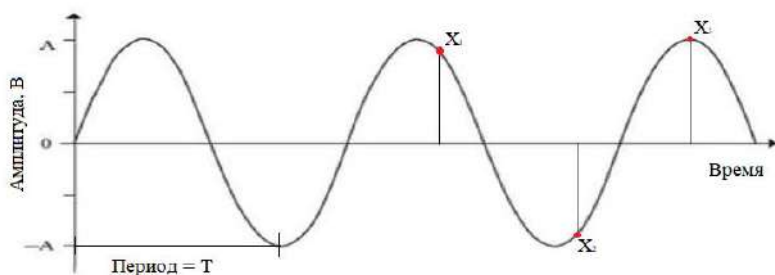


Рис.1 – Схематичное изображение взятия отсчётов

Для восстановления полученных отсчётов в правильном порядке используем следующий алгоритм: временные метки делят по модулю на период восстановленного сигнала (1), затем сортируют от меньшего к большему, сохраняя при этом взаимосвязь между значениями

амплитуды и остатками от деления. Таким образом получаем набор координат для построения одного периода восстанавливаемого сигнала.

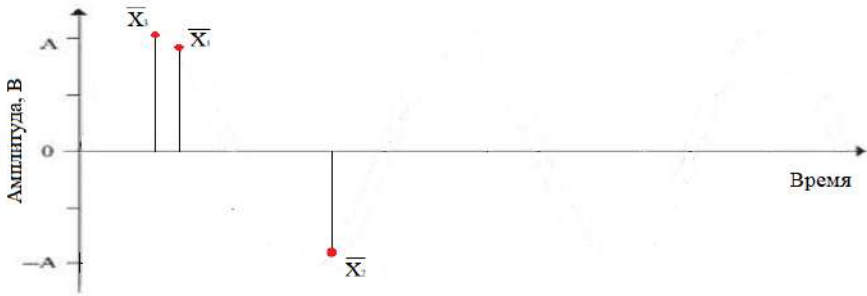


Рис.2 – Отсчёты, по времени разделённые по модулю на период сигнала

Затем получаем функцию через аппроксимацию отсчётов, в качестве аппроксимации используем метод наименьших квадратов.

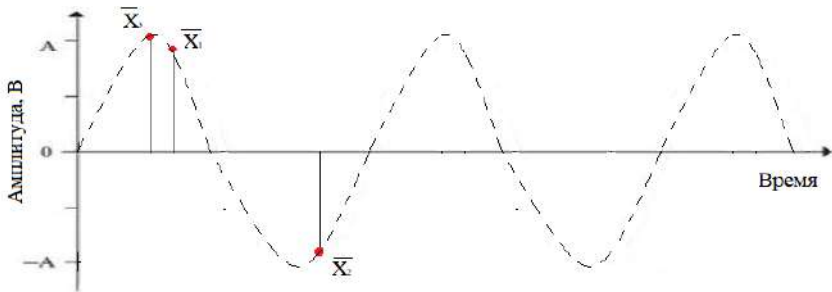


Рис.3 – Пунктирной линией - восстановленный сигнал

Для проверки работоспособности данного метода восстановления сигнала воспользуемся средой Mathcad.

Для модуляции используем функцию следующего вида (2):

$$U(t) = \frac{\sin(\omega \cdot t)}{1} + \frac{\sin(3 \cdot \omega \cdot t)}{9} + \frac{\sin(5 \cdot \omega \cdot t)}{25} \quad (2)$$

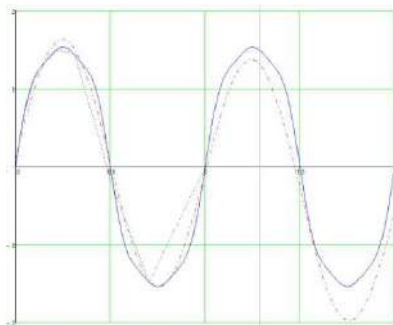


Рис.4 – Реализация метода восстановления в Mathcad

На рис.4 синим цветом показан сгенерированный сигнал по функции (2), красным пунктиром изображён восстановленный сигнал.

Литература:

1. Васюков В.Н. Общая теория связи: Учебник. - Новосибирск: Изд-во НГТУ, 2017. – 580 с.
2. Васюков В.Н. Теория электрической связи: Учебник - Новосибирск, Изд-во НГТУ, 2005. – 392
3. Зюко А.Г. Теория передачи сигналов: Учебник для вузов / А. Г. Зюко, Д. Д. Кловский, М. В. Назаров, Л. М. Финк. – М.: Радио и связь, 1986. – 304 с.

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: НЕЙРОСЕТЕВЫЕ АТАКИ И ЗАЩИТА

А.В. Галашина

**Новосибирский государственный технический университет,
г. Новосибирск, nastya_galashina12@mail.ru**

Научный руководитель: Харюткина С.А., ст. преп.

В статье рассматривается проблема adversarial attacks на нейросетевые модели, которые представляют собой серьезную угрозу информационной безопасности в эпоху искусственного интеллекта. Описываются различные типы adversarial attacks, включая white-box attacks, black-box attacks, poisoning attacks, evasion attacks, targted attacks и untargted attacks, а также анализируют их потенциальные угрозы и последствия для безопасности и надежности систем, основанных на искусственном интеллекте. В статье также представлен обзор современных методов защиты от adversarial attacks.

The article deals with the problem of adversarial attacks on neural network models, which pose a serious threat to information security in the era of artificial intelligence. Various types of adversarial attacks are described, white-box attacks, black-box attacks, poisoning attacks, evasion attacks, targeted attacks, and their potential threats and consequences for the security and reliability of systems based on artificial intelligence are analyzed. The article also provides an overview of modern methods of protection against adversarial attacks.

За последние десятилетия ИИ совершил значительный скачок от теоретических концепций до практических решений. Современные нейросетевые модели демонстрируют выдающиеся результаты в различных областях, но их безопасность остается актуальной проблемой. Одной из актуальных угроз для систем машинного обучения является adversarial attacks (состязательные атаки). Это метод намеренного искажения входных данных, незаметного для человека, которое заставляет модель выдавать ошибочные предсказания. Например, добавление невидимых примесей к изображению может обмануть систему распознавания лиц и открыть доступ к конфиденциальным данным [1].

Существует несколько типов состязательных атак, которые различаются по подходам и целям. В white-box атаках злоумышленник имеет полный доступ к архитектуре и параметрам модели, что позволяет ему проектировать высокоточные искажения. Напротив, в black-box атаках злоумышленник вынужден действовать на основе метода проб и ошибок, так как доступ к внутренним данным модели отсутствует [1]. Некоторые атаки, такие как targeted, направлены на достижение конкретного результата, например, классификацию изображения как определённого объекта. Другие, такие как untargeted, нацелены на то, чтобы модель просто выдала неверный результат, независимо от его характера [2]. Кроме того, атаки могут проводиться на этапе обучения, когда злоумышленник намеренно искажает обучающие данные, чтобы создать предвзятость в модели, либо на этапе эксплуатации, вводя специально созданные данные, которые вынуждают систему работать некорректно. Атаки на обучение (poisoning attacks) искажают данные, чтобы повлиять на модель, что может привести к ошибкам в критических системах. Атаки на инференс (evasion attacks) обманывают модель с помощью специально созданных данных, делая её непригодной для использования в критических ситуациях [1].

Рисунок 1 показывает классификацию атак, разделённых на атаки по расположению и по уровню знаний, включая примеры, такие как отравление данных, атаки обхода, белые и чёрные ящики [1].

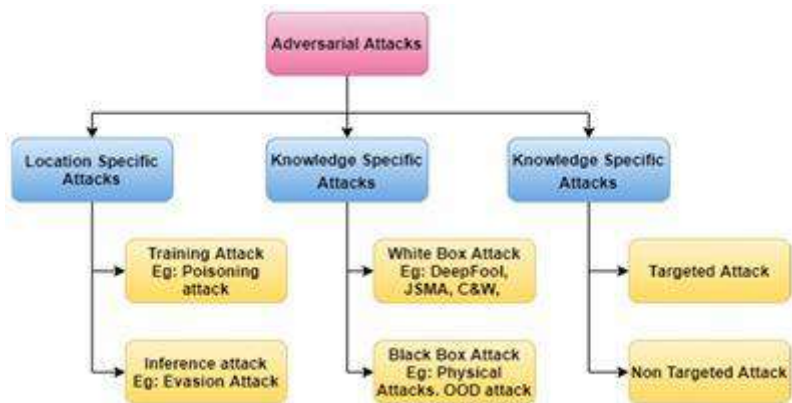


Рис. 1 – классификация атак

Последствия состязательных атак могут быть крайне серьезными и варьируются в зависимости от области применения. Например, модель, подвергшаяся атаке, может ошибочно классифицировать объекты, что приводит к сбоям в системах распознавания лиц, ошибочным медицинским диагнозам или мошенничеству в финансовых системах [2]. Важной угрозой является обход систем безопасности: атаки позволяют злоумышленникам преодолевать биометрическую аутентификацию или получать доступ к защищённой информации. Ошибки в работе ИИ, вызванные такими атаками, подрывают доверие пользователей, что может негативно сказаться на репутации компаний и привести к юридическим или финансовым потерям. Вдобавок к этому, атакованные модели требуют больше вычислительных ресурсов, что снижает их производительность [1].

Для противодействия состязательным атакам разработаны различные методы. Один из наиболее эффективных способов защиты — это обучение модели на adversarial примерах. Этот метод включает в себя создание и добавление к обучающему набору данных примеров, которые были модифицированы с помощью adversarial атак. Таким образом, модель учится распознавать и правильно обрабатывать такие примеры. Существуют различные техники генерации adversarial примеров, такие как Fast Gradient Sign Method (FGSM) и Projected Gradient Descent (PGD) [2]. Аугментация данных — это метод искусственного увеличения обучающего набора, путем создания новых образцов данных, путем преобразования существующих. Эти преобразования (вроде поворотов, масштабирования, изменения

яркости и т.д.) создают вариативность данных, что делает модель более устойчивой к незначительным изменениям во входных данных и снижает вероятность обмана (атак). Регуляризация — это способ предотвратить переобучение модели, чтобы она лучше обобщала на новые данные. Методы, вроде L2-регуляризации или Dropout, добиваются этого, делая модель менее чувствительной к мелким изменениям в обучающих данных. Некоторые архитектуры моделей специально создаются, чтобы противостоять атакам с использованием adversarial examples. Например, использование ансамблей моделей повышает устойчивость, потому что разные модели реагируют на эти примеры по-разному. Также существуют нейронные сети, такие как Capsule Networks, которые лучше противостоят подобным атакам. [1]

В статье рассмотрены основные типы adversarial attacks, их особенности и влияние на безопасность и надёжность систем ИИ. Несмотря на достижения в разработке методов защиты, проблема остаётся значимой и требует дальнейших исследований и совершенствования подходов для обеспечения устойчивости и безопасности технологий, играющих важную роль в современном мире.

Литература:

1. Patil S.; Varadarajan V.; Walimbe D.; Gulechha S.; Shenoy S.; Raina A.; Kotecha K. Improving the Robustness of AI-Based Malware Detection Using Adversarial Machine Learning // Algorithms. – 2021, <https://doi.org/10.3390/a14100297> (дата обращения 17.11.2024).
2. Akhtar N., Mian A. Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey // Journal Of Latex Class Files. – 2017, arxiv.org/pdf/1801.00553 (дата обращения 17.11.2024).

**РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ
ВИДЕОНАБЛЮДЕНИЯ И ПРОГНОЗА АВАРИЙ
НА ИНДИВИДУАЛЬНЫХ ТЕПЛОВЫХ ПУНКТАХ**

Г.Г. Перевалов, Н.А. Блинников

**Новосибирский государственный технический университет,
г. Новосибирск, perevalov.2019@corp.nstu.ru,
blinnikov.2019@stud.nstu.ru**

Научный руководитель: Яковина И.Н., к.т.н., доцент

В данной работе рассматривается решение задачи разработки интеллектуальной системы видеонаблюдения и прогноза аварий на индивидуальных тепловых пунктах, позволяющей на основе потока входных данных и использования современных технологий искусственного интеллекта в режиме мониторинга осуществлять динамический анализ и оценку ситуации и формирования выходных индикаторов о состоянии объекта наблюдения.

This work addresses the development of an intelligent video surveillance and accident prediction system for individual heating stations. The proposed solution enables dynamic analysis and assessment of the situation in real-time monitoring by utilizing incoming data streams and modern artificial intelligence technologies. The system generates output indicators reflecting the state of the observed object, providing actionable insights for improved safety and reliability.

Целью работы является разработка моделей и методов автоматического анализа потока входных данных и видеопотока, позволяющего выявлять аномалии в работе индивидуальных тепловых пунктах (ИТП) и осуществлять прогноз аварийных ситуаций на основе интеграции разнородных данных. На основе анализа существующих подходов к решению аналогичных задач и проведенных экспериментов были разработана коллекция моделей с использованием языка программирования Python, методов повышения точности и устойчивости алгоритмов. Разработка таких решений включает поиск решений для нескольких задач: построение модулей интеллектуального видеонаблюдения для автоматического обнаружения опасных ситуаций, создание моделей для прогнозирования аварий на основе имеющихся данных, а также интеграцию результатов обработки видеопотоков и данных, поступающих с датчиков для повышения аналитических возможностей.

В данном исследовании создание модели проводилось на данных с ИТП Детского сада 171 города Новосибирска за период с 5 октября 2021

года по 27 апреля 2022 года с интервалом измерений каждые 5 минут. Данные о режиме работы поступают через JSON-запросы, где каждый объект представляет собой набор четырех измерений с временной меткой и значениями различных параметров, измеренных в градусах Цельсия: Tsup, Tback, SP_sup, SP_back. Для улучшения качества работы модели была проведена дополнительная подготовка данных, включающая добавление признаков: «часовые циклы», «день недели», «день месяца» и приведение параметров к единому масштабу для ускорения обучения и повышения стабильности работы модели. Прогноз осуществлялся для того же ИТП на данных за период с 5 октября 2022 года по 27 апреля 2023 года с тем же интервалом в 5 минут. Полученная модель позволила с высокой точностью рассчитать температуру выходящей воды в контуре отопления на основе исторических данных и дополнительных признаков. На Рис. 1 представлены действительные (синим цветом) и прогнозные (красным цветом) значения температуры выходной воды. Нейронные сети с долгой краткосрочной памятью Long Short-Term Memory (LSTM) [1, 2]. В выбранном варианте архитектуры LSTM принимает последовательности временных рядов длиной 288 шагов (один день с интервалом в 5 минут), где каждый шаг описан 7 признаками (основные показатели и дополнительные параметры). Эти данные подаются в LSTM-слой из 100 нейронов с сигмоидальной функцией активации для анализа временных зависимостей, краткосрочных и долгосрочных паттернов. Для предотвращения переобучения используется Dropout с коэффициентом 0.2, который случайно отключает 20% нейронов во время обучения. Выходные данные передаются в полносвязный слой (Dense) с 4 нейронами, каждый из которых предсказывает одну из целевых переменных: Tback, Tsup, SP_back, SP_sup. Модель обучается, минимизируя среднеквадратичную ошибку (mse) с помощью оптимизатора Adam, что позволяет эффективно находить зависимости в сложных временных данных. Архитектура нацелена на точное предсказание целевых переменных, учитывая временные циклы и тренды.

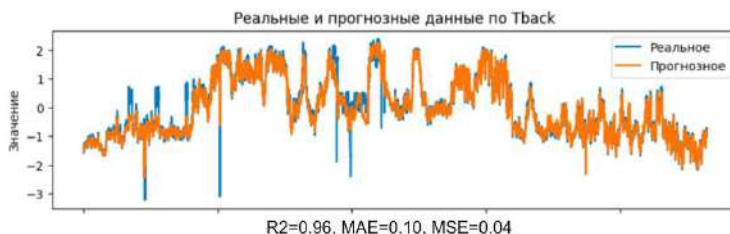


Рис. 1 - Пример работы полученной модели на параметрах Tback.

Машинное зрение является важным дополнением к интеллектуальным системам ИТП. Алгоритмы анализируют видеопотоки в реальном времени и автоматически определяют присутствие человека в опасных зонах. Нейронные сети способны в реальном времени обрабатывать видеопоток, идентифицируя движение и распознавая подозрительные действия. В сочетании с системой оповещения такие меры существенно повышают уровень безопасности [3].

Для решения задачи распознавания объектов в видеопотоке (Рис. 2) использовалась модель нейронной сети YOLOv8 (You Only Look Once, версия 8), которая представляет собой усовершенствованную архитектуру для решения задачи обнаружения объектов, работающую в режиме реального времени. Принцип её действия основан на одноэтапном подходе, при котором входное изображение анализируется полностью за один проход нейронной сети, без разделения на этапы вычленения признаков и классификации. Модель делит изображение на сетку и одновременно предсказывает координаты ограничивающих рамок объектов, их классы и уверенность в распознавании. YOLOv8 отличается улучшенной точностью и скоростью работы благодаря обновлённым слоям нейросети, более эффективным методам обработки признаков и оптимизации алгоритмов. Эти преимущества позволяют использовать ее для решения задач, требующих быстрой и точной обработки больших потоков данных [5]. С учётом необходимости оптимизации вычислительных ресурсов и снижения нагрузки на оборудование было принято решение провести дополнительное дообучение модели. Этот процесс был направлен на адаптацию YOLOv8 для более точного распознавания специфичных объектов для решаемой задачи.



Рис. 2 - Области определения чрезвычайных ситуаций и проникновений в ИТП

Модель способна обрабатывать входные данные, поступающие с любых видеокамер, благодаря предварительной нормализации изображения до стандартизированного разрешения 640x640 пикселей. На этапе предобработки кадры автоматически масштабируются до указанного формата, что позволяет обеспечить совместимость с архитектурой модели и унифицировать процесс обработки данных. После выполнения анализа результат может быть преобразован обратно в исходное разрешение либо адаптирован под иные требуемые параметры, обеспечивая гибкость применения модели в различных сценариях видеонаблюдения. При наличии локальных данных, полученных с места предполагаемого использования, модель на основе YOLOv8 может быть дообучена для оптимизации её работы. Уменьшение числа классов в процессе дообучения позволяет сосредоточить вычислительные ресурсы на распознавании целевых объектов, что способствует повышению точности их определения. Кроме того, адаптация модели к специфике задачи не только улучшает её производительность, но и снижает требования к вычислительным мощностям, облегчая использование в системах с ограниченными ресурсами.

Интеграция предиктивной аналитики на основе LSTM-сетей с системой машинного зрения создаёт комплексную систему мониторинга индивидуальных тепловых пунктов (ИТП). Предиктивная аналитика позволяет прогнозировать аварийные ситуации по техническим параметрам, тогда как машинное зрение обеспечивает визуальный контроль опасных зон, обнаруживая присутствие человека и физические аномалии, такие как дым или пар [4]. Синергетическая обработка данных, поступающих из различных источников, позволяет техническому персоналу получать более комплексную и

детализированную информацию о состоянии индивидуальных тепловых пунктов (ИТП). Это способствует оперативному и обоснованному принятию решений, снижая риск аварийных ситуаций и минимизируя возможный ущерб. Дополнительно автоматизация процессов мониторинга и системы автоматического оповещения уменьшают нагрузку на персонал, что не только оптимизирует трудозатраты, но и повышает общую эффективность эксплуатации оборудования.

Литература:

1. *Бадика Е.М., Зырянов Д.А.* (2021). Предсказание временных рядов при помощи нейронных сетей [Электронный доступ] // Инновации Наука Образование. – URL: <https://elibrary.ru/item.asp?id=47990551> (дата обращения: 7.11.2024).
2. *Андреева Д.О., Тимофеева Д.К., Астапов В.Н.* Аналитический обзор использования искусственных нейронных сетей в промышленности [Электронный доступ] // Международный студенческий научный вестник. – URL: <https://elibrary.ru/item.asp?id=50017258> (дата обращения: 7.11.2024).
3. Enhancing Safety at Industrial and Manufacturing Facilities with Intelligent Behavioral Recognition Video Analytics [Электронный доступ] / Viisights. – URL: <https://www.viisights.com> (дата обращения: 7.11.2024).
4. *Digital Barriers.* Why Real-Time Video Can Upgrade Health and Safety in Manufacturing and Logistics [Электронный доступ]. – URL: <https://digitalbarriers.com> (дата обращения: 7.11.2024).
5. Ultralytics. YOLOv8 Documentation [Электронный доступ] // Ultralytics. YOLO Docs. – URL: <https://docs.ultralytics.com/models/yolov8/> (дата обращения: 7.11.2024).

РАЗРАБОТКА МОДУЛЯ ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ МОБИЛЬНОГО РОБОТА

Н.Е. Галета

Новосибирский государственный технический университет,

г. Новосибирск, nikitagaleta2002@yandex.ru

Научный руководитель: Яковина И.Н. к.т.н.

В статье рассматривается задача разработки в симуляторе CoppeliaSim модуля технического зрения для мобильного робота с системой автопилота, предназначенного для автономного передвижения в изменяющихся условиях окружающей среды. Использование симулятора для тестирования и

оптимизации алгоритмов позволяет воссоздавать различные сценарии и формировать технологически выверенное решение поставленной задачи.

The paper deals with the task of developing a vision module in the CoppeliaSim simulator for a mobile robot with an autopilot system designed for autonomous movement in changing environmental conditions. The use of simulator for testing and optimization of algorithms allows to recreate various scenarios and form a technologically verified solution to the task.

Введение

Системы технического зрения для мобильных роботов являются неотъемлемой частью современных технологий автономной навигации и управления. Использование различных сенсоров для восприятия окружающей среды позволяет роботам ориентироваться и взаимодействовать с объектами в реальном времени. Модули технического зрения находят применение в автономных транспортных средствах, роботах-помощниках, системах мониторинга и инспекции, а также в области автоматизации на промышленных предприятиях [1].

Разработка модуля технического зрения для мобильного робота с автопилотом выполнялась в симуляторе CoppeliaSim. В качестве сенсора для восприятия окружающей среды была использована виртуальная камера (vis_sensor), а также робот Pioneer3DX, который является одной из стандартных платформ для тестирования алгоритмов автономной навигации [2]. В рамках работы реализована обработка видеопотока с камеры и разработаны алгоритмы для обнаружения объектов и препятствий. Вся разработка осуществлялась с использованием языка программирования Python, который позволяет эффективно интегрировать различные библиотеки для обработки изображений и управления роботом.

В статье представлены основные этапы работы: создание модели системы технического зрения, ее тестирование в симуляторе и исследование работоспособности алгоритмов в различных сценариях автономного движения.

Тестовая сцена для проверки работоспособности модуля технического зрения включает мобильный робот Pioneer3DX, оснащённый системами восприятия окружающей среды (см. рисунок 1).

Строение мобильного робота

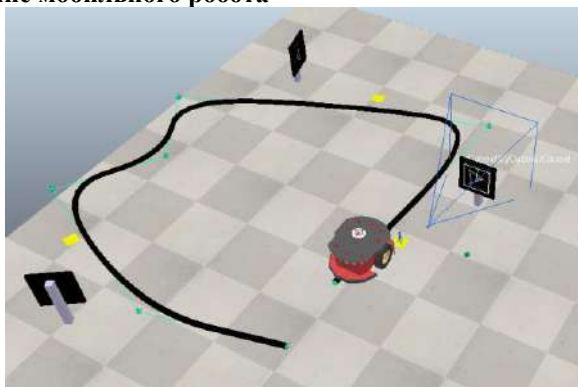


Рис. 1 – Иерархия сцены в симуляционной среде.

Сцена содержит следующие объекты:

Мобильный робот Pioneer3DX, который движется по заранее подготовленной дорожной разметке.

Дорожные знаки пяти типов: ромб, треугольник с указанием направления вправо, треугольник с указанием направления влево, квадрат и круг. Эти знаки служат для тестирования способности системы к их классификации.

Камеры, направленные вниз, предназначенные для распознавания дорожной разметки и отслеживания меток, сигнализирующих о необходимости считывания знаков.

Камера, направленная вперёд, обеспечивающая детекцию и классификацию дорожных знаков, а также анализ дорожной разметки перед роботом.

Эта сцена позволяет проверить работу разрабатываемого модуля в условиях, приближенных к реальным, включая взаимодействие с дорожной инфраструктурой и выполнение корректировок движения на основе визуальных данных.

Функции и работа модуля машинного зрения для мобильного робота

Модуль машинного зрения для мобильного робота принимает на вход изображения, а также информацию о характеристиках моторов и камер. Основные функции разработанного модуля: предобработка изображений, обнаружение и распознавание дорожных знаков, а также оценка и планирование траектории движения робота на основе данных, поступающих с камер.

Обнаружение и распознавание объектов

Методы обнаружения и распознавания объектов, такие как SIFT (Scale-Invariant Feature Transform) и SURF (Speeded-Up Robust Features), демонстрируют высокую устойчивость к изменениям масштаба и ориентации объектов, обеспечивая надежное обнаружение ключевых точек (см.Таблицу 1). Современные подходы на основе сверточных нейронных сетей (CNN) значительно превосходят классические методы по производительности, обучаемости и масштабируемости, что делает их предпочтительным выбором для большинства приложений [3]. При анализе методов и сравнение был выбран подход использования свёрточной сети (CNN). Для её реализации использовалась библиотека программирования Tensorflow [4].

Таблица 1 – Сравнение методов

Критерии	SIFT	SURF	CNN
Обнаружение ключевых точек	+	+	+
Устойчивость к изменениям объектов	+	+	+
Производительность	-	-	+
Обучаемость и масштабируемость	-	-	+

Программная реализация:

Для обработки информации, поступающей с симуляционной модели мобильного робота, использовался язык программирования Python в сочетании с симуляционной средой CoppeliaSim. Python обеспечивал взаимодействие между компонентами системы и реализацию алгоритмов машинного зрения, включая предобработку изображений, распознавание объектов и анализ движения. Симулятор CoppeliaSim позволял моделировать работу робота в виртуальной среде, обеспечивая интеграцию данных с виртуальных сенсоров, таких как камеры и другие устройства, установленные на роботе.

Тестирование подтвердило надежность решения: робот демонстрировал устойчивость к шуму и изменению освещения. Нейронная сеть показала точность классификации знаков более 95%. Лучше всего распознавались геометрически четкие знаки, такие как «Круг» и «Ромб», тогда как знаки «Правый треугольник» и «Левый треугольник» распознавались хуже из-за схожести контуров. Для каждого типа дорожного знака проводилось по 130 экспериментов.

Робот успешно следовал траектории, а применение П-регулятора обеспечивало плавное движение и эффективную корректировку скорости колес. Это подтвердило адаптивность и стабильность разработанной системы.

Заключение:

Модуль машинного зрения был протестирован в симуляционной среде CoppeliaSim в различных сценариях, включая движение по прямой линии, распознавание дорожных знаков и корректировку траектории. В рамках тестов робот стартовал с начальной позиции, двигался по трассе, корректировал траекторию с помощью PI-регулятора, ориентируясь на изменения яркости линии, и останавливался при обнаружении желтой метки для классификации дорожного знака. После распознавания знака робот выполнял маневры в соответствии с заданным сценарием.

Литература:

1. Иванов В.А., Петров А.С., Кузнецов М.В. Обзор современных систем технического зрения, применяемых в транспортной отрасли [Электронный доступ] // Современные наукоемкие технологии. – 2023. – № 6. – С. 12-20. – URL: <https://top-technologies.ru/ru/article/view?id=40150> (дата обращения: 07.11.2024).
2. Coppelia Robotics [Электронный ресурс] / CoppeliaSim Python API. URL: <https://www.coppeliarobotics.com/helpFiles/en/remotedApiFunctionsPython.html> (дата обращения: 05.11.2024).
3. *Fausett, L. Fundamentals of Neural Networks: Architectures, Algorithms and Applications.* Prentice Hall.
4. Documentation [Электронный ресурс] / TensorFlow. – URL: <https://www.tensorflow.org/docs> (дата обращения: 03.11.2024).

РАЗРАБОТКА ПЛАТФОРМЫ ДЛЯ ПРОДАЖИ ПРОГРАММНОГО КОДА

С.В. Панарин

**Новосибирский государственный технический университет,
г. Новосибирск, qraycts@gmail.com
Научный руководитель: Новицкая Ю.В., ст. преп.**

Рассмотрена разработанная платформа, функционирующая в мессенджере Telegram, направленная на решение проблем традиционных фриланс-платформ. Платформа предлагает заказчикам готовые решения, что упрощает процесс

выбора и ускоряет внедрение программного обеспечения. Использование мессенджера Telegram обеспечивает удобство и высокую доступность, а также эффективное управление взаимодействием через чаты и базу данных, обеспечивая высокий уровень персонализации.

The platform developed in the work, functioning in the Telegram messenger, is aimed at solving the problems of traditional freelance platforms. The platform offers customers ready-made solutions, which simplifies the selection process and speeds up the implementation of software. Using Telegram messenger provides convenience and high accessibility, as well as effective management of interaction through chats and a database, ensuring a high level of personalization.

В последние годы наблюдается значительный рост потребности в высокоэффективных, масштабируемых и гибких решениях в сфере информационных технологий [1]. Для разработки таких решений компании могут привлекать как собственных разработчиков, так и независимых исполнителей через платформы фриланса. Однако у платформ фриланса существуют определенные недостатки, такие как отсутствие контроля над исполнителями и возможные нарушения сроков выполнения работ.

В работе представлена созданная платформа, которая направлена на устранение указанных недостатков. Платформа предназначена для продажи кода и функционирует в мессенджере Telegram, что обусловлено несколькими ключевыми факторами. Во-первых, мессенджер Telegram обладает широкой популярностью и активным пользовательским сообществом, что обеспечивает быструю доступность и удобство использования для заказчиков и исполнителей. Во-вторых, реализация платформы на базе мессенджера Telegram позволяет значительно ускорить процесс разработки и внедрения решения, так как мессенджер уже имеет готовую инфраструктуру для общения, что сокращает время на разработку и интеграцию.[2]

Платформа включает чат, в который поступают сформированные заявки, а также чат-бот, обеспечивающий основное взаимодействие между заказчиками и исполнителями. Кроме того, платформа содержит базу данных, в которой хранится вся информация о пользователях и исполнителях, что позволяет эффективно управлять взаимодействием, отслеживать историю заявок и обеспечивать высокий уровень персонализации. Заявка в контексте данной платформы представляет собой код, предлагаемый исполнителем.

Рисунок 1 демонстрирует архитектуру разработанной платформы, иллюстрируя ключевые компоненты, такие как чат-бот и взаимодействие через чат, а также базу данных для управления заявками и пользователями.

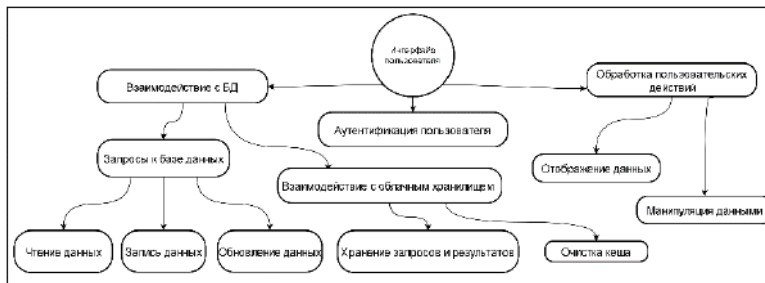


Рис. 1 – архитектура разработанной платформы

Основным отличием разработанной платформы от других платформ фриланса является то, что заказчик имеет возможность выбирать не исполнителей, а уже готовые решения. Это существенно упрощает процесс выбора и сокращает время нахождения необходимого программного обеспечения или кода, так как заказчику не нужно проходить этапы подбора исполнителей и согласования условий. Вместо этого он может сразу приступить к использованию готовых решений, что повышает эффективность рабочего процесса и снижает риски, связанные с исполнением заказов. Главным преимуществом такой концепции является возможность моментально находить необходимые решения и избегать всех недостатков, описанных выше.

В процессе разработки платформы использовались современные технологии и библиотеки, обеспечивающие её функциональность и производительность. Для взаимодействия с мессенджером Telegram использована библиотека telebot, которая позволяет интегрировать чат-бота и управлять заявками пользователей. Для работы с базой данных применяется СУБД MySQL. Для обработки платежей используется система YooKassa, которая позволяет интегрировать безопасную оплату через платформу. S3_client обеспечивает взаимодействие с облачным хранилищем для хранения данных и файлов.

Таким образом, разработанная в мессенджере Telegram платформа не только устраняет основные недостатки традиционных фриланс-платформ, но и предоставляет пользователям удобный и быстрый доступ к готовым решениям, что способствует улучшению процессов в сфере информационных технологий.

Литература:

1. *Сысоев А.* Итоги 2023 года для ИТ-рынка России: главные тренды и прогнозы на 2024 [Электронный ресурс] // РБК Компании. –

URL:

<https://companies.rbc.ru/news/4F89PZdKP1/itogi-2023-goda-dlya-it-ryinka-russii-glavnyie-trendyi-i-prognozyi-na-2024/> (дата обращения: 10.11.2024).

2. Kolesnikov, N. 55+ Telegram Statistics You Need to Know in November 2024 – Revenue, Usage, and More [Электронный ресурс] // business2community. – URL:

<https://www.business2community.com/statistics-pages/telegram-statistics> (дата обращения: 10.11.2024).

ВЫЯВЛЕНИЕ ВЗАИМОСВЯЗЕЙ МЕЖДУ ПОВЕДЕНИЕМ ПОЛЬЗОВАТЕЛЕЙ ИНТЕРНЕТ-МАГАЗИНОВ И ФУНКЦИЯМИ СОЦИАЛЬНЫХ СЕТЕЙ

Л.И. Чуркина

**Новосибирский государственный технический университет,
г. Новосибирск, churkina.2019@stud.nstu.ru**

Научный руководитель: Бакаев М.А., к.т.н., доцент

Исследование анализирует влияние функций социальных сетей на поведение пользователей C2C платформы электронной коммерции. Рассматриваются функции, такие как количество подписчиков, подписок, лайков и использование списка желаний, и их влияние на активность и покупки. Данные собраны с ведущей платформы модной одежды. Результаты могут помочь в разработке маркетинговых стратегий.

The study analyzes the impact of social media features on the user behavior of a C2C e-commerce platform. It examines features such as the number of followers, subscriptions, likes, and wishlist usage and their impact on activity and purchases. The data is collected from a leading fashion platform. The results can help in developing marketing strategies.

Понимание поведения онлайн-покупателей имеет решающее значение для оптимизации пользовательского опыта и улучшения продаж. Исследователи неоднократно пытались выявить аспекты, влияющие на покупательское поведение пользователей во время онлайн-покупок как на сайтах электронной коммерции [1, 2], так и в социальных сетях [3, 4]. Однако влияние типичных для социальных сетей функций на покупательское поведение пользователей в онлайн-магазинах остается недостаточно изученным.

Целью данного исследования является анализ влияния функций социальных сетей на платформе электронной коммерции C2C (потребитель-потребитель) на покупательское поведение пользователей платформы.

Результаты могут стать основой для дальнейших исследований и разработки моделей, направленных на улучшение функционала платформы и удовлетворение потребностей пользователей.

Набор данных для этого исследования включает пользовательские данные успешного интернет-магазина модной одежды C2C с более чем 10 миллионами зарегистрированных пользователей [5]. Платформы C2C, позволяют людям покупать товары и продавать их. На платформе предусмотрены функции социальных сетей, например, пользователи могут подписываться друг на друга, ставить лайки товарам, добавлять их в листы желаний. Магазин впервые был запущен в Европе примерно в 2009 году, а затем распространился по всему миру.

Набор данных включает 98913 строк и 24 столбца, среди которых хэш идентификатора пользователя, демографические данные, информация о количестве подписчиков и подписок пользователя, о количестве понравившихся товаров, товаров в списке желаний, купленных товаров, проданных товаров, загруженных непроданных товаров. Некоторые столбцы повторяют информацию из других столбцов и представляют ее в другом виде.

В исследовании была выделена группа активных пользователей. В нее попадают те, кто взаимодействовал с товарами: загрузили, но еще не продали товар, купили, продали, лайкнули, добавили товар в список желаний. Также в отдельную группу были выделены пользователи, которые использовали функции социального взаимодействия, то есть подписывались на учетные записи или получали подписчиков.

Гипотезы исследования формулируются следующим образом.

H1: Большая часть пользователей – активные пользователи (загрузили, но еще не продали товар, купили, продали, лайкнули, добавили товар в список желаний).

H2: Использование функций социального взаимодействия (подписки на учетные записи пользователей или получение подписчиков) положительно связано с уровнем активности пользователей.

H3: Количество подписок пользователей положительно связано с объемом покупок пользователей.

H4: Использование списка желаний положительно связано с объемом покупок пользователей.

H5: Число лайков пользователей положительно связано с их объемом покупок.

H6: Количество подписчиков продавца положительно связано с его объемами продаж.

Для проверки первой гипотезы была использована дескриптивная статистика. Для проверки остальных гипотез исследования был

использован регрессионный анализ. Регрессионный анализ позволяет проверять статистические гипотезы о значимости влияния независимых переменных на зависимую переменную. Поэтому для каждой гипотезы предварительно были выделены независимые и зависимые переменные. Использование регрессионного анализа требует выполнения некоторых условий, а именно:

1. Остатки (разницы между наблюдаемыми и предсказанными значениями) должны иметь нормальное распределение.
2. Между остатками не должно быть корреляции, значит, остатки должны быть независимыми.

Для проверки зависимости остатков друг от друга в модели регрессии используется тест Дарбина-Уотсона (DW). Чтобы исключить корреляцию между остатками значение DW должно находиться в диапазоне от 1,5 до 2,5.

Результаты показали, что только 21.9% пользователей являются активными, большинство пользователей никак не взаимодействовали с товарами, таким образом H1 не подтверждается.

Таблица 1 – Влияние функций социального взаимодействия на уровень активности пользователей

Предиктор	B	R ²	p-значение	DW
Кол-во подписчиков	4.163	0.652	0.000	1.941
Кол-во подписок	-1.822			

В Таблице 1 исследуется влияние функций социального взаимодействия на уровень активности пользователей. В (Коэффициент регрессии) показывает влияние предиктора на зависимую переменную. В данном случае в качестве функций социального взаимодействия выступают количество подписчиков и количество подписок. Параметр $B = 4.163$ для количества подписчиков означает, что с получением одного нового подписчика, активность пользователей увеличивается на 4.163 единицы в среднем, а для количества подписок $B = -1.822$ означает, что при увеличении числа подписок на одну единицу целевая переменная уменьшается на 1.822. Это может означать, что большое количество подписок отрицательно связано с целевой переменной, что может свидетельствовать о других типах поведения пользователя.

R-значение равно 0.000, что говорит о том, что уровень социального взаимодействия статистически значим (влияет на количество покупок) на уровне 0.05. $R^2 = 0.652$ означает, что модель объясняет 65.2% вариации зависимой переменной (уровень активности) через независимую переменную (уровень социального взаимодействия). Это

довольно высокая доля, что указывает на сильную связь между социальным взаимодействием и активностью.

Значение Дарбина-Уотсона находится в допустимом диапазоне, то есть между остатками нет корреляции.

Таблица 2 – Влияние количества подписок на объем покупок пользователей

Предиктор	B	R²	p-значение	DW
Кол-во подписок	0.099	0.013	0.000	1.878

Таблица 3 – Влияние списка желаний на объем покупок

Предиктор	B	R²	p-значение	DW
Объем списка желаний	0.039	0.152	0.000	1.882

Таблица 4 – Влияние количества лайков на объем покупок

Предиктор	B	R²	p-значение	DW
Кол-во лайков	0.016	0.052	0.000	1.9

Таблица 5 – Влияние количества подписчиков продавца на его объем продаж

Предиктор	B	R²	p-значение	DW
Кол-во подписчиков	0.733	0.583	0.000	1.769

По таблицам 2, 3, 4 и 5 можно сказать, что связи в них статически значимы, значение Дарбина-Уотсона находится в допустимом диапазоне, а также:

между количеством подписок и объемом покупок, объясненная вариация очень мала ($R^2 = 0.013$), что делает связь между ними слабой (Таблица 2).

15.2% вариации объема покупок объясняется объемом списка желаний (Таблица 3).

влияние числа лайков на объем покупок пользователей слабое (Таблица 4).

58.3% вариации объема продаж объясняется количеством подписчиков продавца, что указывает на высокую значимость этого фактора (Таблица 5).

Таким образом, среди всех пользователей активными являются лишь 21.9%. Гипотезы H3, H4 и H5 показали слабую связь между зависимыми и независимыми переменными. Количество подписчиков оказало сильное влияние и на объем покупок (H2), и на объем продаж (H6).

Ограничения и будущие исследования:

1. Низкие значения R^2 в гипотезах H3, H4 и H5 показывают, что большая часть вариации объема покупок не может быть объяснена только изученными предикторами, такими как подписки, списки желаний и лайки.

2. Анализ не учитывает такие факторы, как качество продукта, ценовая политика, сезонность и маркетинговые активности, которые могут значительно влиять на объем покупок.

3. В данном исследовании не учитывались демографические характеристики пользователей (возраст, пол, регион). Эти параметры будут учтены в дальнейших исследованиях.

Литература:

1. *Koehn D., Lessmann S., Schaal M.* Predicting online shopping behaviour from clickstream data using deep learning // *Expert Systems with Applications*. – 2020. – Vol. 150. – P. 113342.

2. *Safara F.* A computational model to predict consumer behaviour during COVID-19 pandemic // *Computational Economics*. – 2022. – Vol. 59. – N. 4. – P. 1525-1538.

3. *McClure C., Seock Y.K.* The role of involvement: Investigating the effect of brand's social media pages on consumer purchase intention // *Journal of retailing and consumer services*. – 2020. – Vol. 53. – P. 101975.

4. *Liao S.H., Widowati R., Hsieh Y.C.* Investigating online social media users' behaviors for social commerce recommendations // *Technology in Society*. – 2021. – Vol. 66. – P. 101655.

5. E-commerce - Users of a French C2C fashion store [Electronic resource] // Kaggle. – 2019. – URL: <https://www.kaggle.com/datasets/jmmvutu/ecommerce-users-of-a-french-c2c-fashion-store/data> (accessed 01.11.2024).

РАЗРАБОТКА СИСТЕМЫ АВТОМАТИЗАЦИИ БИЗНЕС-ПРОЦЕССОВ

М.О. Денисов, И.Н. Яковина

Новосибирский государственный технический университет,
г. Новосибирск, denisov.00779@gmail.com
Научный руководитель: Яковина И.Н. к.т.н., доц. каф. ВТ.

В работе рассматривается создание ERP системы, которая направлена на использование аналитических инструментов для обеспечения гибкого управления персоналом, финансовыми потоками и проектами. Материал работы показывает средства разработки основ системы, использование библиотек для языка программирования python. Описана структура онлайн-сервиса с примерами базовых функций и функциональных особенностей. Также приведены результаты тестирования скорости работы приложения и сравнения ранее использованного технического решения.

The paper discusses the creation of ERP system, which aims to use analytical tools to provide flexible management of staff, financial flows and projects. The material of the work shows the tools for developing the fundamentals of the system, the use of libraries for the programming language python. The structure of the online service with examples of main functions and functional features is described. The results of testing the speed of the application and comparison of the previously used technical solution are also given.

Актуальность разработки

В нынешней ситуации, когда большинство зарубежных компаний не могут стабильно предоставлять доступ к своему программному обеспечению, многим российским компаниям приходится искать аналоги популярных решений. Например, компании Microsoft и Amazon закрыли доступ к своим облачным сервисам для российских компаний. Зачастую такими решениями становятся собственные разработки, специфика которых направлена на нужды конкретной компании заказчика. Таким образом, разработка софта, который закроет проблемы ручной монотонной работы путем автоматизации бизнес-процессов, становится целью для текущей работы.

Такая задача является актуальной для быстро расширяющейся инженерной компании, где с ростом числа сотрудников, увеличивается и количество работы по: аналитике данных, учету доходов и расходов, подсчет статистики работы для каждого сотрудника, а также предоставление статистики, как индивидуально для каждого сотрудника, так и для общей статистики компании для вышестоящих руководителей. Старые инструменты корпоративного учета и аналитики стали работать заметно медленнее. Как говорилось выше, зарубежные

компании прекращают свою облачную деятельность, значит, прежнее рабочее решение перестают быть актуальным. Для решения ряда этих задач была начата внутренняя разработка собственного программного обеспечения, которое в удобной форме могло бы предоставить полный функционал получения и редактирования данных, хранящихся на удаленном сервере.

Постановка задачи, используемые методы и средства

Разрабатываемый сервис по функционалу наиболее схож с системой планирования и учета корпоративных ресурсов (ERP). Ее основными функциями также выступают: управление финансами, персоналом, цепочками производства и гибким внедрением отдельных микросервисов. В ходе выполнения задачи были решены основные задачи по подбору главных инструментов, создание базы данных, разработка серверной части приложения, основных алгоритмов автоматизации для сбора и предоставления хранимых данных, тестирование сотрудниками компании и анализ полученных данных в ходе тестирования, подведение промежуточных итогов.

Оборудование: Главным средством для хранения базы данных и всех файлов для запуска приложений в виде докер контейнеров является удаленный сервер под управлением операционной системы Linux Ubuntu. Он настроен для работы как в локальной, так и в глобальной сети, для предоставления нужной страницы по запросу [4].

Программное обеспечение: для обеспечения реализации всех возможностей системы были выбраны следующие инструменты: язык программирования Python, pandas для работы с таблицами данных [1], Streamlit для визуального отображения [2], фреймворк Django для написания backend составляющей работы проекта, в том числе работы с базой данных, СУБД PostgreSQL для хранения данных [5], xlswriter для промежуточного сохранения данных в excel документ [3]. Окончательный вид всей системы наиболее схож с микросервисной архитектурой, где каждый сервис имеет свое API для взаимодействия.

Особенности разработки

Разработка схемы БД и сбор данных. Для аналитики, считывания, записи и создания данных была выбрана реляционная СУБД PostgreSQL. В числе основных преимуществ перед остальными системами можно выделить производительность и масштабирование, что очень важно для расширяющийся в реальном времени компании. Также, система управления обладает многоверсионным управлением конкурентным доступом, что означает изоляцию транзакций, поскольку чтение, запись и обновление данных может производиться разными пользователями одновременно. Поскольку доступ к данным и к работе с

ними может понадобиться в любой момент времени плюс данной СУБД станет непрерывность действий, что предоставит доступ к данным даже во время сбоя БД.

Разработка алгоритмов автоматизации и работы с БД. Одной из самых важных функций при работе с сервисом является предобработка полученных из БД данных и предоставление их пользователю в виде таблиц, графиков или метрик. Для загрузки и предобработки используется библиотека Pandas [1]. В backend части приложения осуществлено подключение к БД. Для получения нужных данных из таблиц реализованы запросы. Далее данные проходят через редактирование. Pandas имеет функции редактирования таблиц аналогичные Power BI [6].

Веб-интерфейс сервиса. Для визуализации информации, чтения и редактирования было выбрана библиотека streamlit, которая позволяет получить доступ ко всем ресурсам посредством обычной интернет-страницы [2]. Достоинством этого решения является возможность использования множества готовых визуальных шаблонов, таких как:

- Текстовые элементы: названия, заголовки, подписи, подзаголовки, поля для ошибок и предупреждений;
- Элементы отображения данных: статичные и редактируемые таблицы, метрики и графики;
- Элементы взаимодействия в виде кнопок, которые позволяют активировать заданные функции, скачивать данные, переходить по ссылкам и обращаться к различным сервисам;
- Элементы фильтрации: переключение, выбор одной опции, выбор множества опций, выбор даты или промежутка дат.

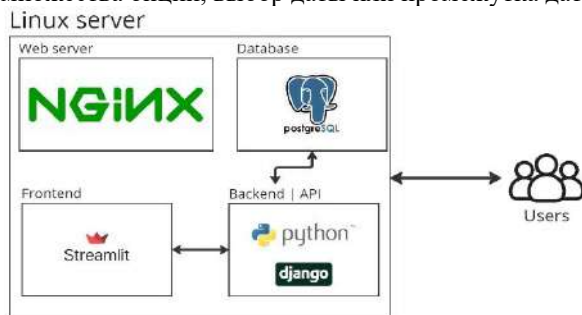


Рис. 1 – Архитектура сервиса управления ресурсами.

Функциональность разработанного приложения

Из всех вышеперечисленных компонентов разработки складывается общая архитектура системы. Она состоит из множества сервисов, каждый из которых использует для своей работы базу данных, имеет собственную backend составляющую для запроса и редактирования данных, API для получения и отправки запросов остальным сервисам. В таблице 1 представлены точечные сравнения времени работы при использовании разрабатываемого сервиса и прошлого основного решения Power BI. Для доступа пользователей ко всем функциям системы разработан общий веб-интерфейс.

Микросервисная архитектура системы позволяет:

- Хранить и взаимодействовать со всеми данными посредством СУБД PostgreSQL;
- Перенаправлять запросы пользователей для доступа к сервисам и распределять нагрузку между ними с помощью Nginx;
- Предоставлять по запросу с опцией обработки данных информацию из базы данных используя Python django;
- Выводить всю необходимую информацию в предобработанном виде, добавлять, удалять и обновлять данные с помощью веб-интерфейса, разработанного в streamlit;
- Скачивать сформированные excel документы.

Таблица 1 – Временные характеристики работы

Характеристики сравнения	Power BI	Текущая система
Время загрузки, отображения данных	~30 секунд	~0.1 секунда
Время обновления данных	>5 минут	~1 секунда
Время редактирования данных табличных	>2 минут	~0.1 секунда

Заключение

Таким образом, в ходе разработки онлайн системы автоматизации бизнес-процессов были решены следующие задачи: подбор инструментов, подходящих под нужды проекта; разработка базы данных и сбор данных для ее наполнения; реализация основных алгоритмов работы данными и их обработка. В ходе тестирования была разработана

универсальная форма визуализации информации, а также реализованы инструменты для ее фильтрации и отображения требуемого среза данных. Также, были проведены измерения ключевых метрик при работе системы и сравнения их с ранее используемым в организации решением. Общая архитектура разработанной системы состоит из множества сервисов. Это позволяет гибко настраивать набор функций для разных групп пользователей.

Литература:

1. Python data analysis library (2023) [Electronic resource] / Pandas. – URL: <https://pandas.pydata.org/docs/> (accessed: 01.11.24)
2. Open-source Python library (2023) [Electronic resource] / Streamlit. – URL: <https://docs.streamlit.io> (accessed: 02.11.24).
3. Python module (2024) [Electronic resource] / XlsxWriter. <https://xlsxwriter.readthedocs.io/index.html> (accessed: 03.11.24).
4. Ubuntu (2024) [Electronic resource] / Ubuntu Server. – URL: <https://ubuntu.com/server> (accessed: 05.11.24).
5. Transaction Isolation (2024) [Electronic resource] / PostgreSQL 16. – URL: <https://www.postgresql.org/docs/current/transaction-iso.html> (accessed: 05.11.24).
6. Power BI (2024) [Electronic resource] / Microsoft. – URL: <https://www.microsoft.com/ru-ru/power-platform/products/power-bi?market=ru#Interoperability> (accessed: 04.11.24).

РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ СОТРУДНИКОВ КАДРОВОЙ СЛУЖБЫ ОРГАНИЗАЦИИ

С.М. Хабаров, И.Н. Яковина

**Новосибирский государственный технический университет,
г. Новосибирск, Khabarov.semen@mail.ru,**

Научный руководитель: Яковина И.Н., к.т.н., доцент

В данной работе рассматривается разработка информационной системы, направленной на интеграцию и группировку данных из различных источников для проведения аналитики и обеспечения оперативного доступа к информации в реальном времени. Целью системы является повышение скорости и продуктивности работы сотрудников кадровой службы компании в условиях импортозамещения.

This work focuses on the development of an information system aimed at integrating and aggregating data from various sources to enable analytics and provide real-time access to information. The system's goal is to enhance the speed

and productivity of HR department employees within the framework of import substitution. The project centers on creating a unified platform that offers convenient tools for data analysis and visualization, facilitating effective decision-making and optimizing personnel management processes.

С введением международных санкций и уходом крупных западных компаний, таких как Google, с российского рынка, многие бизнес-процессы оказались под угрозой. В частности, кадровой службы, которые долгое время использовали западные сервисы и приложения для работы с данными, анализа и построения дашбордов и графиков для анализа, столкнулись с серьезными трудностями. Инструменты, такие как Google Sheets, Google Analytics, Looker Studio, которые раньше обеспечивали хранение, обработку и визуализацию данных, теперь либо недоступны, либо имеют ограниченный функционал. Эти ограничения усложняют процесс работы с данными и снижает продуктивность.

Ранее у сотрудников кадровой службы данные хранились в различных Google Таблицах, что значительно усложняло доступ к информации и тормозило рабочие процессы. Каждая таблица содержала отдельные аспекты статистики и для получения полной картины требовалось постоянно переключаться между несколькими источниками. Это не только снижало эффективность работы, но и увеличивало время на обработку данных и принятие решений. Дополнительные задачи, такие как подготовка аналитических отчетов или создание персонализированных дашбордов, требовали отдельной выгрузки данных из баз данных, что добавляло еще один этап в процесс, замедляя его и увеличивая вероятность ошибок. Всё это создавало значительные препятствия для быстрой и качественной работы отдела кадровой службы.

Для решения этой проблемы было принято решение создать единую веб-платформу, которая обеспечивала бы объединяла данные с различных ресурсов, таких как базы данных, локальные хранилища и другие системы, и интегрировала бы их в одном интерфейсе. Это позволило бы не только упростить доступ к информации, но и реализовать возможность визуализации данных в реальном времени, построения графиков и аналитических отчетов, что в свою очередь повысит скорость принятия решений и продуктивность работы сотрудников кадрового отдела. Кроме того, при проектировании платформы предусмотрено ее масштабирование, что позволит адаптировать систему для других направлений бизнеса.

Минусы отечественных программ мы рассмотрим на примере Яндекс Таблиц и Яндекс Datalends(Дашборды) по сравнению с инструментами Google можно обозначить следующим образом:

Минусы Яндекс Таблиц:

1. Функционал: Яндекс Таблицы уступают Google Sheets по количеству функций.
2. В Яндекс Таблицах возможности автоматизации и интеграции более ограничены.
3. Совместная работа: несмотря на наличие совместного редактирования, в Google Sheets оно более стабильно и часто удобнее в использовании, особенно при одновременной работе большого числа пользователей.
4. Расширения и аддоны: В Google Sheets есть множество аддонов, которые можно установить для расширения функциональности. Яндекс Таблицы не имеют аналогичного набора инструментов.
5. Ограничение по объему данных: Google Sheets позволяет хранить больше данных в одном документе по сравнению с Яндекс Таблицами.

Минусы Яндекс Datalends:

1. Визуализация данных: Google Lookerstudio предоставляет более мощные и гибкие инструменты для создания визуализаций.
2. Интеграции с Google-сервисами: Google Lookerstudio легко интегрируется с множеством источников данных от Google, таких как Google Analytics, Google Ads, BigQuery, а также с внешними базами данных.
3. Гибкость настроек: В Google Lookerstudio более широкий спектр настроек дашбордов и виджетов, включая использование параметров и фильтров.
4. Глобальная поддержка и локализация: Google Lookerstudio является глобальным продуктом с обширной поддержкой и документацией на разных языках.
5. Скорость работы: В Google Data Studio данные могут обновляться быстрее благодаря высокопроизводительным серверам и более развитой инфраструктуре.

Стек технологий:

1. Backend:
Python – основной язык программирования.
Django REST Framework – фреймворк для создания веб-приложений и API [1].
MySQL – реляционная база данных для хранения и управления данными.
2. Обработка данных:
Pandas – библиотека для работы с данными в формате DataFrame.
NumPy – библиотека для работы с массивами и матрицами данных, а также для научных вычислений.

SQLAlchemy – ORM для Python, позволяющая эффективно работать с базой данных через Python-код.

OS (модуль Python) – стандартная библиотека Python для взаимодействия с операционной системой, предоставляющая функции для работы с файловой системой, процессами и окружением.

3. Визуализация данных:

Matplotlib – библиотека для создания статических, анимационных и интерактивных графиков и диаграмм.

Seaborn – расширение для Matplotlib.

Plotly – библиотека для создания интерактивных графиков и дашбордов.

Bokeh – инструмент для создания интерактивных визуализаций, работающих в браузере.

Этот стек технологий обеспечит надежную и масштабируемую платформу с возможностью эффективной обработки и визуализации данных для отдела кадровой службы.

Когда фронтенд запрашивает данные, Django REST Framework [2] на бэкенде принимает запрос и с помощью SQLAlchemy выполняет SQL-запросы к базе данных MySQL. Полученные данные обрабатываются Pandas и NumPy — для анализа, очистки и подготовки к визуализации. Далее с помощью Matplotlib, Seaborn создаются статические графики, а Plotly и Bokeh позволяют реализовать интерактивные визуализации, которые передаются обратно на фронтенд через API и отображаются в браузере пользователя.

Страница статистики, представленная на рисунке 1, разделена на несколько направлений, с оформленными секциями для графиков и таблиц. Такой подход позволяет легко анализировать и просматривать ключевую статистику в удобном и структурированном формате.

Заключение:

Информационная система, представляется крайне важной задачей в условиях импортозамещения. Уход западных компаний создал необходимость в эффективных отечественных решениях, и примеры, такие как Яндекс, подтверждают, что российские аналоги пока не способны полностью заменить зарубежные продукты.

Таким образом, создание подобной платформы не только решает проблему текущего дефицита инструментов для работы с данными, но и способствует снижению зависимости от иностранных сервисов, поддерживая процесс импортозамещения и улучшая внутреннюю эффективность компании.

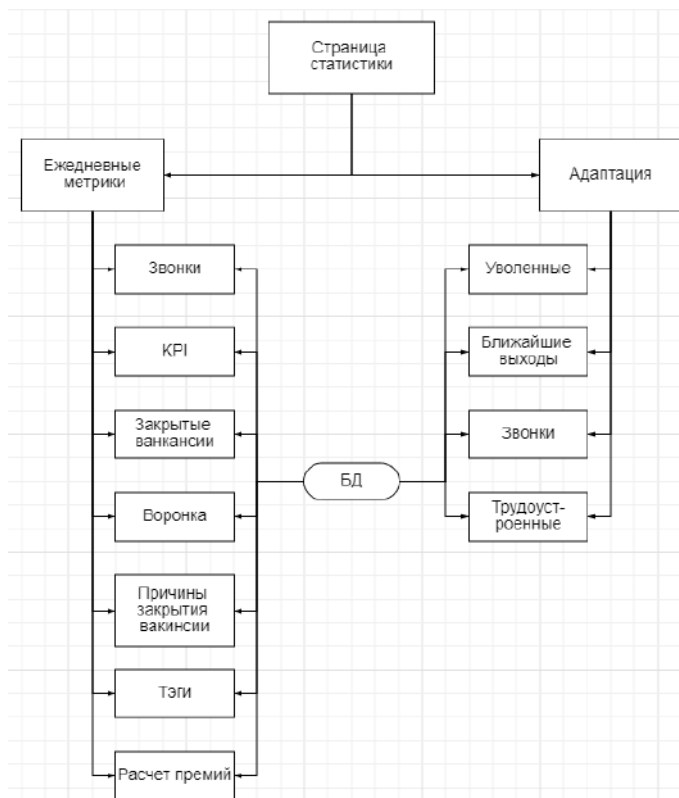


Рис. 1 – структура страницы статистики

Литература:

1. Для чего используется Django REST Framework [Электронный ресурс] // NIC, 2024. – URL: https://www.nic.ru/help/dlya-chego-ispol6zuetsya-django-rest-framework_11684.html (дата обращения: 01.09.2024).
2. Создание современных веб-API с Django и Django REST Framework: Полное руководство [Электронный ресурс] // DEV Community, 2024. – URL: <https://www.developerchronicles.com/django-rest-framework-api-guide> (дата обращения: 15.09.2024).

РАЗРАБОТКА ИНСТРУМЕНТОВ АВТОМАТИЧЕСКОЙ МИГРАЦИИ ИНФОРМАЦИОННЫХ СИСТЕМ С ORACLE 11G НА СУБД POSTGRES PRO

И.Г. Масленников

**Новосибирский государственный технический университет,
г. Новосибирск, maslennikov.2020@stud.nstu.ru**

Научный руководитель: Стасышин В.М., к.т.н., доцент

В данной работе рассматриваются программные инструменты, позволяющие осуществить автоматическую миграцию информационных систем с СУБД Oracle 11g на Postgres Pro. Электронная система НГТУ, включающая в себя множество различных компонентов, в том числе и СУБД Oracle 11g, требует современного подхода к управлению данными и их миграции. Переход на более современные и гибкие решения, такие как Postgres Pro, открывает новые возможности для оптимизации и повышения эффективности работы с данными.

This paper examines software tools that enable automatic migration of information systems from Oracle 11g DBMS to Postgres Pro. The NSTU electronic system, which includes many different components, including Oracle 11g DBMS, requires a modern approach to data management and migration. The transition to more modern and flexible solutions, such as Postgres Pro, opens up new opportunities for optimizing and increasing the efficiency of working with data.

Электронная система НГТУ включает в себя множество компонентов, среди которых особую роль играет база данных Oracle 11g. Эта СУБД обеспечивает хранение и управление данными для различных приложений, используемых в образовательном процессе. Однако, несмотря на свою надежность и функциональность, Oracle 11g сталкивается с рядом ограничений, таких как высокие затраты на лицензирование, сложность масштабирования и уход компании Oracle Corporation с российского рынка.

Переход на более современную и гибкую СУБД, такую как Postgres Pro, позволяет значительно улучшить управление данными и оптимизировать процессы. Важным аспектом этого перехода является разработка инструментов автоматической миграции, таких как oga2pg, которые упрощают и ускоряют процесс переноса данных и объектов базы данных.

Эти инструменты обеспечивают высокую степень автоматизации, минимизируя риск ошибок и потерь данных. В данной работе последовательно будут рассмотрены подходы и решения, направленные на успешную миграцию информационных систем с Oracle 11g на

Postgres Pro, а также анализ преимуществ, которые это предоставляет для образовательной среды НГТУ.

I. Цели и задачи

Основная проблема, которую помогает решить данная задача – переход госорганов и ключевых предприятий РФ на отечественные софт и оборудование до 2030 года. Постановление об этом подписал премьер-министр РФ Михаил Мишустин, документ есть в распоряжении ТАСС.

Речь идет о субъектах критической информационной инфраструктуры, к которым относятся государственные органы и предприятия или индивидуальные предприниматели, использующие информационные системы в критически важных отраслях, таких как здравоохранение, наука, транспорт, связь, энергетика и другие.

Помимо глобальных целей, стоит задача провести целостное исследование, включая теоретическую и практическую части, по используемому СУБД. Разработать скрипты, которые в последствие могут быть применены в различных структурах и областях российской отрасли. Миграция также обусловлена необходимостью обеспечить независимость от иностранных вендоров и адаптацию к изменениям на рынке. Основные задачи включают перенос всех данных и объектов базы данных, настройку инфраструктуры, а также проверку корректности работы системы на новом оборудовании.

II. Этапы миграции

Этапы миграции обеспечивают успешный переход системы НГТУ на новую СУБД. Каждый этап выполняет определенные задачи и обеспечивает плавность переноса данных и функциональности, с минимальными рисками для целостности данных и доступности системы.

1. Анализ и оценка исходной системы

На текущем этапе проводится комплексное исследование текущей системы на базе Oracle 11g для выявления всех аспектов, которые повлияют на процесс миграции. Основные задачи включают:

Инвентаризация объектов базы данных, где фиксируются все таблицы, индексы, представления, триггеры, последовательности и другие объекты базы данных необходимые для переноса.

В анализе хранимых процедур и триггеров проводится оценка PL/SQL-кода для выявления сложных запросов и процедур, которые могут потребовать модификации для совместимости с Postgres Pro.

Определение объема данных: оценивается размер данных для расчета необходимой мощности инфраструктуры Postgres Pro и продолжительности миграции.

Этап идентификации бизнес-логики, связанной с БД заключается в выявлении зависимости между приложениями и базой данных, чтобы минимизировать влияние на работу системы во время миграции.

2. Подготовка инфраструктуры

После завершения анализа нужно подготовить систему к переносу на новое оборудование или конфигурацию, если это необходимо:

Настройка серверов и сетевого оборудования – немаловажный этап в выборе и настройке серверов, которые будут использоваться для Postgres Pro, включая CPU, RAM и дискового пространства, необходимых для производительности.

Далее непосредственно установка Postgres Pro и настройка необходимых параметров, таких как размер буферов, лимиты подключений и конфигурация WAL-журналов для записи транзакций.

Подготовка резервного хранилища подразумевает создание системы резервного копирования для обеспечения возможности восстановления данных при ошибках или сбоях.

3. Перенос схемы и данных

Основной этап включает в себя процесс миграции объектов и данных из Oracle 11g в Postgres Pro с использованием инструментов автоматизации, таких как ora2pg:

Конвертация схемы базы данных, то есть выполнение с помощью ora2pg преобразования структур таблиц, индексов, представлений и других объектов, адаптируя их под формат Postgres Pro.

После конвертации схемы ora2pg переносит сами данные. Этот процесс может включать настройку для ускорения переноса (например, временное отключение индексов).

Далее каждая таблица сверяется с исходными данными в Oracle 11g для подтверждения корректности.

4. Проверка и тестирование

По завершении переноса данных необходимо тщательно протестировать работоспособность системы и соответствие результатов. Проводится функциональное тестирование, в котором проверяются ключевые функции приложений, взаимодействующих с базой данных, чтобы удостовериться, что данные загружаются и обрабатываются корректно. Производственное тестирование выполняет нагрузки для оценки скорости обработки запросов и других критичных операций. И один из заключительных этапов – тестирование отказоустойчивости, где проводится имитация отказов для проверки работоспособности системы в условиях нагрузки и способности восстановления после ошибок.

5. Оптимизация и настройка

На конечном шаге происходит детальная оптимизация работы базы данных под Postgres Pro для достижения оптимальной производительности и надежности. Анализируются и оптимизируются SQL-запросы, которые могут требовать модификаций для более эффективной работы в Postgres Pro. Настраиваются индексы и параметры конфигурации, позволяющие значительно ускорить обработку данных.

III. Инструмент Ora2Pg

Очень хорошо показала себя open-source утилита с говорящим названием ora2pg. Утилита ora2pg подключается к исходной СУБД Oracle и считывает метаданные из словаря СУБД (Oracle Dictionary), на выходе утилита формирует SQL-скрипты создание таблиц, индексов, ограничений целостности и т.д. в синтаксисе Postgres Pro (он отличается от синтаксиса создания объектов в СУБД Oracle).

ora2pg — это мощный инструмент с открытым исходным кодом, предназначенный для автоматизации миграции данных и схем из Oracle в Postgres. Он существенно упрощает и ускоряет процесс перехода, позволяя конвертировать структуры базы данных и переносить данные с минимальными ручными изменениями.

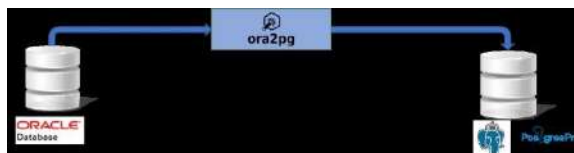


Рис. 1 – Схема работы инструмента Ora2Pg

Таким образом миграция с Oracle 11g на Postgres Pro в рамках информационных систем НГТУ представляет собой стратегически важное решение, направленное на улучшение гибкости, оптимизацию затрат и повышение автономности системы. Проведенный анализ показал, что использование специализированных инструментов автоматической миграции, таких как ora2pg, позволяет значительно упростить и ускорить процесс переноса данных, минимизируя риск потерь и ошибок.

Литература:

1. Поддержка пакетов в стиле Oracle в СУБД Postgres Pro 15 Enterprise [Электронный ресурс] // Habr. – URL: <https://habr.com/ru/companies/postgrespro/articles/726204/> (дата обращения: 29.10.2024).

2. Documentation [Электронный ресурс] // Postgres Pro Enterprise. – URL: <https://postgrespro.ru/docs/enterprise/16/> (дата обращения: 29.10.2024).

РАЗРАБОТКА ПРИЛОЖЕНИЯ «СИСТЕМА ПРЕДОСТАВЛЕНИЯ ДОСТУПА ПОЛЬЗОВАТЕЛЯМ К РЕСУРСАМ СУПЕРКОМПЬЮТЕРНЫХ ЦЕНТРОВ»

С.К. Мельников, В.А. Шахрай

Новосибирский государственный технический университет,

г. Новосибирск, s.melnikov.2019@stud.nstu.ru

Научный руководитель: Аврунев О.Е., директор ЦИУ

В данной работе описывается программное решение, позволяющее упростить работу специалистов технических направлений, нуждающихся в использовании систем с высокой производительностью, таких как суперкомпьютерные центры.

This paper describes a software solution designed to streamline the work of professionals in technical fields who require access to high-performance systems, such as supercomputing centers.

Суперкомпьютерные центры пользуются большим спросом в научном сообществе, производящим расчеты на больших данных благодаря тому, что обладают куда большей производительностью чем обычный стационарный ПК.

Однако получить доступ к суперкомпьютерному центру не всегда удастся легко. Сама процедура предоставления доступа является очень сложным процессом. Например, сначала заполняется форма на веб-странице суперкомпьютерного центра, затем печатается и заполняется анкета, которую необходимо подписать после чего отправить в администрацию суперкомпьютерного центра [1-3]. Через некоторое время будет получен ответ, и если он будет одобрительным, то нужно будет выслать публичный SSH-ключ, для осуществления подключения через удаленный доступ по протоколу SSH с возможностью передачи данных по протоколу SFTP. Иные способы взаимодействия через удаленный доступ могут быть существенно ограничены в целях обеспечения информационной безопасности [4].

После подключения у пользователя возникает следующий ряд проблем. Во-первых, для многих специалистов технических направлений, нуждающихся в возможностях СКЦ может быть незнакома операционная система семейства Unix, установленная на

суперкомпьютере. Во-вторых, работа на суперкомпьютере осуществляется через планировщика заданий SLURM, который обычный пользователь в повседневной жизни не использует. В-третьих, разработанные программы могут не использовать параллелизм и, следовательно, не смогут получить максимальную производительность.

Для упрощения работы пользователей и улучшения её качества было выдвинуто решение в виде веб-приложения, позволяющего эффективно взаимодействовать с суперкомпьютерным центром с момента подачи заявки в суперкомпьютерный центр до завершения работы на нём.

В этой работе будут детально рассмотрены компоненты, разрабатываемого приложения, необходимые для достижения поставленных целей.

Проект представляет собой клиент-серверное приложение, состоящее из нескольких отдельных сервисов, обладающих различной функциональностью.

Прежде чем рассматривать каждый сервис более детально, обозначим необходимую и достаточную функциональность разрабатываемого приложения:

- регистрация и авторизация пользователей;
- подача заявки;
- просмотр и редактирование заявки;
- добавление и редактирование команды;
- просмотр доступных СКЦ и их ресурсов;
- подключение к СКЦ;
- загрузка данных на СКЦ и выгрузка результатов оттуда;
- создание воркфлоу;
- просмотр и редактирование воркфлоу;
- оптимизация работы программы пользователей.

После группировки необходимой функциональности можно выделить четыре сервиса, на которые их можно распределить [5]. Укрупненная схема взаимодействия сервисов представлена на рисунке 1.

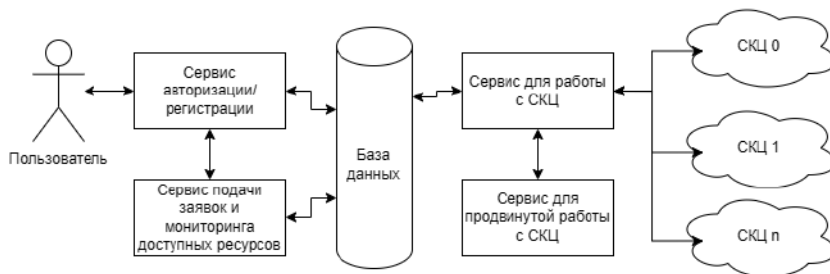


Рис. 7 – Схема взаимодействия сервисов

1. Сервис авторизации и регистрации. Необходим для осуществления идентификации и аутентификации пользователей. Позволяет обеспечить разделение пользователей веб-приложения по группам и ответственности. Вынесен отдельно так как может быть представлен уже существующей службой.

2. Сервис для работы с заявками и мониторинга доступных ресурсов. Осуществляет взаимодействия пользователей с администрацией суперкомпьютерного центра и позволяет устранить сложность с подачей заявок в различные суперкомпьютерные центры благодаря унификации и стандартизации процесса подачи заявки. Также позволяет пользователям подобрать наиболее подходящий для них суперкомпьютерный центр, в зависимости от их запросов и задач.

3. Сервис для работы с суперкомпьютерным центром. Выполняет роль шлюза между всей системой и непосредственно суперкомпьютерным центром. Основные задачи этого сервиса заключаются в обеспечении соединения с суперкомпьютерным центром, получении информации о текущем состоянии СКЦ, обеспечения легкого взаимодействия клиента и СКЦ.

4. Сервис для продвинутой работы с СКЦ. Это расширение для предыдущего модуля, позволяющее модифицировать свой рабочий процесс, делая его наиболее эффективным и, как следствие, увеличивая производительность разрабатываемых программ. Данный сервис занимается анализом разработанного пользователем приложения (отмечает язык программы, выделяет необходимые зависимости и пакеты) и подготавливает рабочее пространство на суперкомпьютере. В свою очередь пользователь отмечает какие программные модули запускать и делает это в виде ориентированного ациклического графа (DAG) [6] с переменными однократного присваивания. После чего это переводится в задачи SLURM, и программа выполняется.

Результатом работы является описание программного комплекса, состоящего из нескольких отдельных сервисов, ориентированных на выполнение задач, вызывающих наибольшие сложности у специалистов при работе с суперкомпьютерными центрами.

Литература:

1. Институт математики [Электронный ресурс]. – URL: <https://math.nsc.ru/sc> (дата обращения: 28.10.2024).
2. Сибирский СКЦ [Электронный ресурс]. – URL: <http://www.sccc.icmmg.nsc.ru/request.html> (дата обращения: 28.10.2024).
3. ИВЦ НГУ [Электронный ресурс]. – URL: <http://nusc.nsu.ru/wiki/doku.php/doc/registration> (дата обращения: 28.10.2024).
4. *Тарасов Н.И.* Веб-лаборатория для суперкомпьютерного многомасштабного моделирования задач напыления / *Н. И. Тарасов, В. О. Подрыга, С. В. Поляков* // Вычислительные методы и программирование: электронный журнал. – URL: <https://en.num-meth.ru/index.php/journal/article/view/1321/1252> (дата обращения: 28.10.2024).
5. Шаблоны архитектуры программного обеспечения [Электронный ресурс] // O'Reilly. – URL: <https://www.oreilly.com/library/view/software-architecture-patterns/9781491971437/ch01.html> (дата обращения: 28.10.2024).
6. Все, что вам нужно знать об Airflow DAGs, ч.1 — Основы и расписания [Электронный ресурс] // Хабр. – URL: <https://habr.com/ru/articles/682384/> (дата обращения: 28.10.2024).

РАЗРАБОТКА СЕРВИСА ПОДАЧИ ЗАЯВОК ДЛЯ ПРИЛОЖЕНИЯ «СИСТЕМА ПРЕДОСТАВЛЕНИЯ ДОСТУПА ПОЛЬЗОВАТЕЛЯМ К РЕСУРСАМ СУПЕРКОМПЬЮТЕРНЫХ ЦЕНТРОВ» И ИНТЕГРАЦИЯ С ИНФОРМАЦИОННОЙ СИСТЕМОЙ НГУ

С.К. Мельников, В.А. Шахрай

**Новосибирский государственный технический университет,
г. Новосибирск, s.melnikov.2019@stud.nstu.ru
Научный руководитель: Аврунев О.Е., директор ЦИУ**

В данной работе описывается разработка сервиса подачи заявок, как части приложения для предоставления доступа к ресурсам суперкомпьютерного центра. А также его интеграция с ИС НГТУ.

This paper describes the development of a service for submitting applications as part of an application for providing access to the resources of a supercomputer center. Additionally, it covers the integration of this service with the NSTU information system.

Суперкомпьютерные центры в зависимости от бюджета их организаций имеют разное техническое оснащение. При поиске суперкомпьютерного центра пользователю необходимо осуществлять выбор наиболее подходящего для него, опираясь на следующие технические характеристики [1-3]:

Общие:

- количество узлов в вычислительной системе;
- объем ПЗУ на каждом узле;
- тип дискового хранилища;
- объем ОЗУ на каждом узле;
- центральный процессор на узлах:
 - тип;
 - количество ядер;
 - количество процессоров;
 - архитектура.

Дополнительные:

- графический процессор:
 - тип;
 - количество ядер;
 - количество процессоров;
 - объем видеопамати;
 - версия ускорителя CUDA.
- нейронный процессор:
 - тип;
 - количество ядер;
 - количество процессоров.
- система хранения данных:
 - объем хранилища;
 - тип хранилища;
 - сетевой интерфейс;
 - тип интерфейса.

Обширное количество параметров в описании технических характеристик суперкомпьютерного центра накладывает дополнительную нагрузку на пользователя, из-за того, что во время выбора придется часто переключаться между интернет-ресурсами СКЦ, для того чтобы сравнить их и определить наиболее подходящий.

Другая проблема заключается в том, что различные суперкомпьютерные центры имеют разные формы для подачи заявок на использование их вычислительных мощностей. В случае отказа в одном суперкомпьютерном центре пользователям приходится каждый раз заполнять новую анкету и повторять процедуру её отправки.

Для решения данной проблемы был разработан веб-сервис для подачи заявок в суперкомпьютерные центры, который станет частью веб-приложения «Система предоставления доступа пользователям к ресурсам суперкомпьютерных центров».

Разработанный сервис обладает следующей функциональностью:

1. Работа с заявкой – создание, редактирование, удаление. Позволяет пользователю создать заявку или её черновик, которую в дальнейшем он сможет отправить на рассмотрение в СКЦ.

2. Сохранение заявок. Позволяет пользователю использовать одну и ту же заявку несколько раз, например, при подаче в другие СКЦ или при повторной подаче в тот же СКЦ.

3. Добавление участников проекта. Так как доступ к СКЦ осуществляется через протокол SSH, то с помощью этой опции можно отмечать тех участников проекта, которым будет необходим доступ к суперкомпьютерному центру, и появится возможность загружать публичные SSH-ключи этих участников.

4. Загрузка резюме руководителя проекта. Администрация СКЦ вправе ознакомиться с тем, кому предоставляет доступ к своим ресурсам.

5. Просмотр технических характеристик суперкомпьютерного центра. Во время подачи заявки отображаются технические характеристики всех доступных СКЦ – рисунок 1, что позволит в удобном формате подобрать для себя наиболее подходящее устройство – рисунок 2.

Фильтры

Укажите необходимые параметры конфигурированной системы, в списке ниже Вы увидите подходящие кластеры

12

CPU

Параметры центрального процессора

Кол-во CPU (ядро)

6

Модель центрального процессора (CPU)

AMD Ryzen 7 5800X

Архитектура CPU

ARM

GPU

Параметры графического процессора

Кол-во GPU (ядро)

4

Модель графического процессора (GPU)

RTX 4080 super

Способность GPU

S.D

NPU

Параметры нейронного процессора

Кол-во NPU (ядро)

2

Модель нейронного процессора (NPU)

PSL Lighting PSL-NPU

Память

Параметры памяти конфигурированной системы

Оперативная память (RAM, TB)

256

Встроенная память (TB)

1024

Тип диска HDD

Рис. 8 – Задание фильтров для вычислительной системы

79

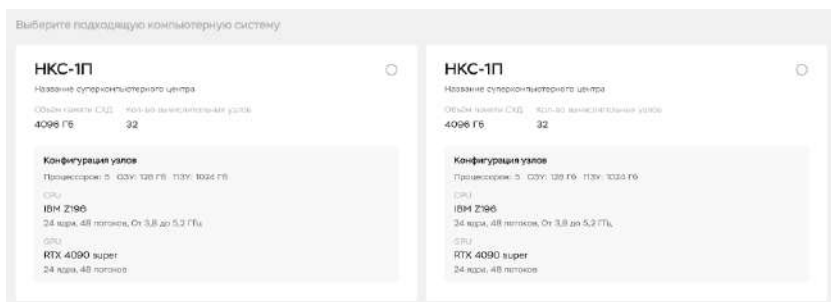


Рис. 9 – Перечень доступных устройств

Для тестирования работоспособности приложения первая его версия была разработана как API для платформы E-Services, являющейся частью цифровой инфраструктуры НГТУ.

Серверная часть сервиса разработана на языке программирования Python с использованием фреймворка Flask. Работа с базой данных осуществляется с помощью ORM SQLAlchemy.

В качестве базы данных была выбрана реляционная СУБД PostgreSQL [4], а в качестве файлового хранилища – MinIO [5]. Данный выбор был сделан на основании того, что обе системы находятся в открытом доступе, распространяются с открытым исходным кодом и просты в освоении.

Клиентская часть написана на языке программирования JavaScript с использованием библиотеки React.

Взаимодействие разработанного сервиса с информационной системой НГТУ осуществляется при помощи публичного API НГТУ. С помощью такого подхода удалось избежать дублирования баз данных, оставив регулирующую роль за ИС НГТУ. Это также позволило протестировать работоспособность сервиса без создания модуля авторизации/регистрации.

Разработанный сервис является важной частью будущего приложения «Система предоставления доступа пользователям к ресурсам суперкомпьютерных центров» и уже сейчас способен помочь многим специалистам сократить время для выбора СКЦ, подходящего под их проекты, и облегчить подачу заявки на него.

Литература:

1. Институт математики: сайт. – URL: <https://math.nsc.ru/sc> (дата обращения: 28.10.2024).
2. Сибирский СКЦ: сайт. URL: <http://www.sscs.icmmg.nsc.ru/request.html> (дата обращения: 28.10.2024).

3. ИВИЦ НГУ: сайт. – URL:
<http://nusc.nsu.ru/wiki/doku.php/doc/registration> (дата обращения: 28.10.2024).
4. PostgreSQL: сайт. – URL:
<https://www.postgresql.org/docs/15/index.html> (дата обращения: 28.10.2024).
5. Minio: сайт. – URL:
<https://min.io/docs/minio/kubernetes/upstream/index.html?ref=docs-redirect> (дата обращения: 28.10.2024).

РАЗРАБОТКА СЦЕНАРИЯ ИЗУЧЕНИЯ МЕТОДОВ КЛАСТЕРИЗАЦИИ СРЕДСТВАМИ ПРОГРАММНОГО ТРЕНАЖЕРА

А.А. Ларионова

**Новосибирский государственный технический университет,
г. Новосибирск, llarionova.an@yandex.ru**

Научный руководитель: Альсова О.К., к.т.н., доцент

Разработан учебный тренажер, предназначенный для изучения иерархических методов кластерного анализа данных. В статье рассматривается сценарий изучения и исследования методов кластерного анализа данных средствами учебного тренажера. Тренажер предоставляет возможность пользователю в рамках концепции no-code (без программирования) исследовать и применять иерархические методы кластерного анализа данных. Для обеспечения эффективного обучения тренажер содержит теоретическую часть, практическую часть с возможностью загружать собственные данные и тестовую часть.

A training simulator designed to study hierarchical methods of cluster data analysis has been developed. The article discusses the scenario of studying and researching methods of cluster data analysis using a training simulator. The simulator provides an opportunity for the user to explore and apply hierarchical methods of cluster data analysis within the framework of the no-code concept (without programming). To ensure effective training, the simulator contains a theoretical part, a practical part with the ability to upload your own data and a test part.

В современном мире объём данных увеличивается с каждым днем. Эти данные необходимо обрабатывать и анализировать. Одной из основных задач анализа данных является кластеризация. Методы кластерного анализа позволяют разбить изучаемую совокупность объектов на группы однородных в некотором смысле объектов, называемых кластерами или классами [1]. Существует множество

алгоритмов кластерного анализа данных. Выбор того или иного алгоритма во многом зависит от специфики прикладной задачи, от структуры исходных данных и априорно обосновать выбор метода не представляется возможным. Это создает дополнительные трудности для специалистов и требует от них комплексных знаний в области кластерного анализа данных. Во многом поэтому с каждым годом потребность в квалифицированных кадрах растет, что мотивирует людей обучаться анализу данных.

В процессе обучения могут использоваться различные подходы, способы представления учебно-методических материалов, программные средства и технологии. Существуют множество онлайн-школ, предоставляющих электронные учебные пособия, видео уроки для самостоятельного образования. Одним из эффективных средств обучения является тренажер. Обучающий тренажер (тренажер) – программный комплекс, ориентированный на обучение человека и отработку им определенных навыков и умений [2]. Данный вид обучения получил большую популярность не только в России, но и за рубежом. Несмотря на наличие множества средств обучения, для изучения методов анализа данных существует сравнительно мало обучающих тренажеров, что делает их разработку особенно актуальной. Следует отметить, что не все действующие тренажеры в полной мере обеспечивают наличие теоретических, практических и тестовых учебных материалов.

В данной статье будет рассмотрен сценарий изучения и исследования иерархических методов кластерного анализа средствами тренажера. Стоит отметить, что в основу тренажера положена концепция по-code, для того чтобы процесс обучения проходил более эффективно.

Учебный тренажер для изучения иерархических методов состоит из 6 частей: теоретические материалы, загрузка данных, стандартизация данных, кластерный анализ, оценка точности и интерпретация результатов, тестирование. На рисунке 1 показан сценарий работы с обучающим тренажером.

При запуске программы пользователю предоставляется выбор, либо ознакомиться с лекционным материалом или же сразу перейти к практике. Теоретическая часть представляет собой четыре лекции, отражающие основной материал по теме «Кластерный анализ». После просмотра лекций, учащемуся предлагается перейти в раздел «Загрузка данных» или пройти тестирование.

В разделе «Загрузка данных» можно загрузить свои данные или же выбрать уже готовый вариант, нажав кнопку «Пример». При первом

знакомстве рекомендуется использовать готовый пример, поскольку этот вариант детально описан в тренажере.

После загрузки данных, их необходимо привести к единой форме, для того чтобы устранить влияние разницы масштабов измерения значений признаков. Для этого в тренажере предусмотрен раздел «Стандартизация данных». В данном разделе учащийся может выбрать, один из предложенных вариантов стандартизации: стандартизация по Z-оценкам, нормализация по методу минимума и максимума, логарифмические преобразования, масштабирование по максимальному абсолютному значению, устойчивое масштабирование. Далее необходимо приступить к кластерному анализу.

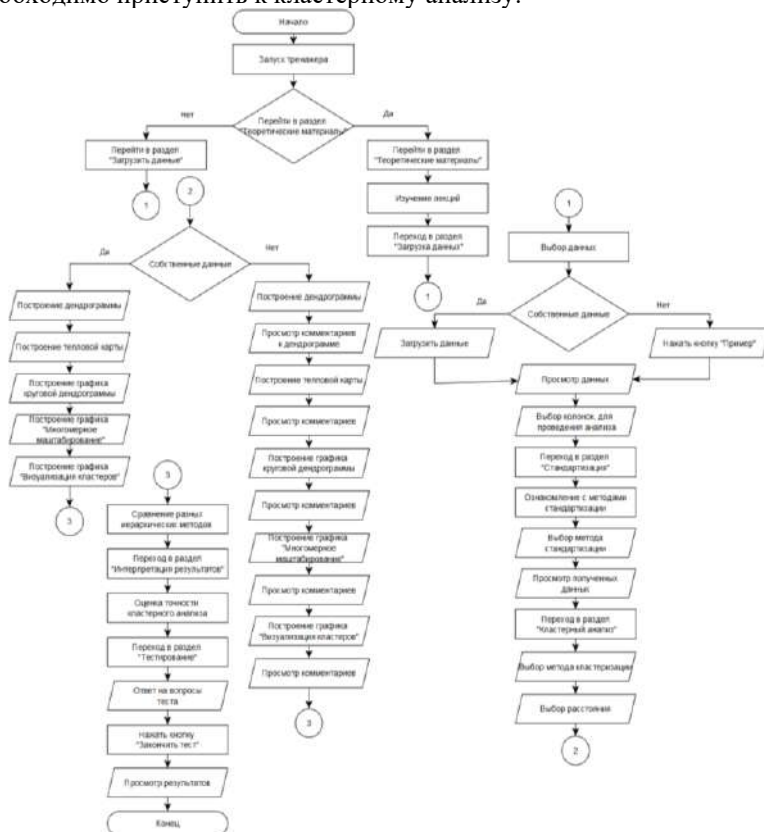


Рис.1 – Сценарий изучения методов кластерного анализа средствами программного тренажера

В разделе «Кластеризация данных», пользователь может наглядно оценить структуру кластеров. Это можно сделать с помощью дендрограммы, тепловой карты и других графиков. Чтобы их построить, необходимо выбрать один из предложенных методов и меру расстояния. В тренажере реализованы такие методы, как метод одиночной связи, метод полной связи, невзвешенное попарное среднее, взвешенное попарное среднее, невзвешенный центроидный метод, взвешенный центроидный метод, метод Уарда, и следующие меры расстояния: евклидово расстояние, взвешенное евклидово расстояние, квадрат евклидова расстояния, расстояние городских кварталов, расстояние Чебышева, степенное расстояние.

Следующим разделом после «Кластерного анализа», является «Оценка точности и интерпретация результатов». В данном разделе, пользователю показывается, как оценивается качество кластеризации.

После проведения кластерного анализа, чтобы проверить как усвоен материал, учащемуся необходимо пройти небольшое тестирование.

В заключение следует отметить, что для эффективного изучения иерархических методов кластерного анализа пользователям рекомендовано следовать последовательности разделов обучающего тренажера. Каждый модуль логически связан с предыдущим и последующим, что позволяет лучше изучать и применять данный материал. Такой подход будет особенно полезен людям, которые только начинают свой путь в анализе данных.

Литература:

1. *Калинина В. Н., Соловьев В. И.* Введение в многомерный статистический анализ: Учебное пособие / ГУУ. – М., 2003. – 66 с.
2. *Казанская О.В. Альсова О.К.* Индуктивные тренажеры в техническом университете // Известия Самарского научного центра Российской академии наук. – 2016. – Т. 18, № 2. – С. 902-906.

ЦИФРОВОЕ СОПРОВОЖДЕНИЕ БОЛЬНЫХ С ХРОНИЧЕСКОЙ СЕРДЕЧНОЙ НЕДОСТАТОЧНОСТЬЮ

А.Ю. Белов, И.Н. Яковина

**Новосибирский государственный технический университет,
г. Новосибирск, andr123b123@gmail.com**

Научный руководитель: Гужов В.И. д.т.н., профессор

В работе представлены некоторые результаты разработки системы анализа и визуализации данных, предназначенной для оценки состояния

пациента и отображения результатов диагностики для их включения в контур поддержки принятия решений в медицинской практике.

The paper presents some results of the development of a data analysis and visualization system designed to assess the patient's condition and display diagnostic results for their inclusion in the decision support circuit in medical practice

Постановка задачи

В ходе разработки системы анализа и визуализации данных на первом этапе был выполнен анализ существующего процесса анализа данных и разработан алгоритм визуализации данных. На втором этапе был спроектирован интерфейс системы визуализации с существующим программным обеспечением, отвечающим за ввод данных. Далее была разработана архитектура приложения, позволяющая использовать библиотеку визуальных объектов для параметрического построения визуального образа в соответствии с характеристиками отображаемых данных.

Сбор и структурирование данных

Источником данных является база данных разрабатываемого регистра хронической сердечной недостаточности, который позволяет для каждого пациента на основании сформированного вектора из 79 параметров строить визуальный диагностический образ по 5 группам данных (см. рисунок 1 и 2). Набор исходных данных и группы диагностических признаков были согласованы с экспертами в области кардиологии и соответствуют актуальным клиническим рекомендациям.



Рис.1 - Информационная модель диагностических данных

Группы параметров могут состоять из подгрупп. Например, для лабораторных тестов – это гормональный анализ крови, общий анализ крови и биохимический анализ крови.

Алгоритм формирования визуальной модели

Необходимость правил расчёта показателей визуального диагностического образа обусловлена потребностью к преобразованию в единый формат представления данных. Для того, чтобы перейти от представления в различных шкалах к единому, используется установленные экспертом уровень «нормы» и правила расчета показателей. В случае, если показатель входит в установленный уровень «нормы», значение X будет равным нулю.



Рис.2 - Пример отображения общей диаграммы состояния пациента

Разберём пример расчёта оценки параметра «ЭхоКГ».

Первый показатель расчёта – фракция выброса (ФВ). Норма от 50 до 70 процентов, допустим, что ФВ для первой госпитализации – 61%, второй – 60%, третьей – 64%. Вывод – показатель входит в норму, соответственно значение приравнивается к нулю.

Следующий показатель – среднее давление в легочной артерии (СДЛА). Норма показателя меньше 25. В качестве примера возьмём, что СДЛА для первой госпитализации – 33, второй – 28, третьей – 43. Во всех трёх госпитализациях показатель превышает установленный уровень нормы и при пересчете показателей будут получены следующие значения:

$$\text{СДЛА}_1 = (33 - 25) / 25 = 0.32;$$

$$\text{СДЛА}_2 = (28 - 25) / 25 = 0.12;$$

$$\text{СДЛА}_3 = (43 - 25) / 25 = 0.72$$

Динамика показателя, рассчитываемого с использованием данной процедуры нормализации хорошо видна на рисунке 2, который отражает результатом сбора и обработки данных трех осмотров одного пациента на одном визуальном образе.

Описание библиотеки визуальных образов

Рассмотрим используемые в библиотеке графики, которые позволят сравнивать между собой различные результаты диагностических данных, характеризующих динамику состояния одного пациента.

Лепестковая диаграмма. Выбор лепестковой диаграммы обусловлен необходимостью просмотра общей ситуации пациента. По лепестковой диаграмме лучше всего отслеживаются изменения в параметрах, какие имеют схожие, какие высокие или низкие значения. Лепестковая диаграмма позволяет видеть, в каких областях произошло улучшение или ухудшение состояния пациента, что способствует более информированному выбору лекарственной терапии и тактики лечения.

Гистограммы. Для групп данных «Лабораторные тесты» и «ЭхоКГ» используются гистограммы. Выбор гистограммы обусловлен необходимостью отслеживать изменения набора показателей и обеспечение их наглядного сравнения в динамике.

Битовые карты. Для групп данных «ЭКГ», «Клиническое состояние» и «Жалобы пациента» используются битовые карты. Они отражают наличие или отсутствие значений в каждом из параметров и позволяют на одном визуальном образе отразить весь спектр состояний.

Алгоритмы обработки и анализа данных

Первый этап обработки данных - сбор необходимой информации из медицинских карт. В процессе необходимо избежать потери или искажения информации. Следующий этап: преобразование данных в единый формат представления. Поскольку различные показатели могут измеряться в различных шкалах, важно привести их к общему виду для дальнейшего анализа. Ключевой этап, на котором осуществляется анализ данных для выявления патологий и отклонений от нормального значения. Используются установленные экспертами уровни “нормы” для различных показателей. Четвертый этап: предоставление визуального отображения реальных показателей пациента. Заключительный этап: на основе анализа данных и выявленных патологий разрабатываются персонализированные рекомендации для пациента. Эти рекомендации помогают специалисту в принятии обоснованных решений по дальнейшему лечению и уходу за пациентом.

Заключение

Разработанный алгоритм позволяет преобразовывать показатели пациента в единый формат, обеспечивая их наглядное отображение и анализ. Использование различных графиков делает визуализацию информативной и удобной для восприятия. Созданный интерфейс

упрощает сбор данных и ускоряет процесс заполнения информации о пациенте, структурируя её по логическим блокам. Интуитивная визуализация позволяет отслеживать динамику состояния пациента, выявлять отклонения от нормы и принимать обоснованные решения для дальнейшей терапии.

Литература:

1. Научный потенциал (2023) // URL: <https://sciencen.org/assets/Kontent/Konferencii/Arhiv-konferencij/KOF-811.pdf> (дата обращения 12.11.2024)
2. Крамм М.Н., Безбородова О.Е., Бодин О.Н., Светлов А.В. Цифровой двойник сердца. – 2021. – №1. – С. 73-84.
3. Цифровые двойники [Электронный ресурс] // ProgramLab: сайт. –URL: <https://pl-llc.ru/lp/mx/pltwins/> (дата обращения 12.11.2024).
4. Суримова В.А., Скородумова Е.А. Создание и интеграция цифрового двойника // Научные технологии в космических исследованиях Земли. – 2022. – Т.14, №4. – С. 54-64.

РАЗРАБОТКА СЕГМЕНТА БАЗЫ МЕДИЦИНСКИХ ЗНАНИЙ ДЛЯ ИНФОРМАЦИОННОЙ СИСТЕМЫ «РЕГИСТР ХСН»

М.В. Рясной, И.Н. Яковина

Новосибирский государственный технический университет,

г. Новосибирск, ryasnoj.2020@stud.nstu.ru

Научный руководитель: Яковина И.Н., к.т.н.

В работе рассмотрено решение задачи разработки сегмента базы медицинских знаний, которая проектируется для использования в контуре информационной системы «Регистр хронической сердечной недостаточности». Организация базы знаний в виде таблиц, правил и алгоритмов позволяет на разных этапах принятия диагностических решений формировать промежуточные результаты анализа данных о пациенте.

The paper considers the solution of the problem of developing a segment of the medical knowledge base, which is designed for use in the outline of the information system “Chronic Heart Failure Registry”. Organization of the knowledge base in the form of tables, rules and algorithms allows to form intermediate results of patient data analysis at different stages of diagnostic decision making.

В настоящее время в среде медицинских информационных систем (МИС) появляется все больше решений для формирования аналитических индикаторов для принятия врачебных решений. В своей основе они используют знания, которые с свою очередь хранятся в

базах знаний МИС [1]. Знания в базах знаний (БЗ) могут быть представлены с помощью различных методов, среди которых наиболее популярными являются таблицы, правил и диагностические алгоритмы.

В работе рассмотрено решение задачи разработки сегмента базы медицинских знаний, которая проектируется для использования в контуре информационной системы «Регистр хронической сердечной недостаточности». БЗ должна включать в себя модули хранения и представления знаний и интерфейс для наполнения информацией. Работоспособность базы знаний должна быть проверена на реальных данных.

В ходе анализа существующих баз медицинских знаний (БМЗ) было выявлено, что существуют различные виды медицинских знаний [2]: таблицы референтных значений; классификационно-диагностические алгоритмы; причинно-следственные паттерны и многие другие. Среди методов представления знания наиболее распространенными в медицинских информационных системах являются [3]: отологии; семантические сети; таблицы решений и правила; логические и вероятностные модели и формализованные клинические рекомендации.

Базы знаний (БЗ) в медицине обеспечивают возможность накопления и систематизации знаний, полученных от специалистов. При использовании этих знаний появляется возможность повысить эффективность процесса диагностики и лечения пациентов. Также применение БЗ облегчает процесс обмена знаниями между специалистами, благодаря систематизированному подходу к хранению и возможности их визуализации.

В ходе решения поставленной задачи был проведен анализ существующих разработок в данной области, спроектирована и разработана база знаний, разработан и реализован интерфейс БЗ, а также выполнено тестирование с использованием реальных данных. Также для проверки корректной работы базы знаний потребовалось разработать и реализовать базу данных с развернутой информацией о пациенте на основании которой формируется набор диагностических индексов в соответствии со структурой диагноза (см. рисунок 1).



Рис.1 - Структура диагноза в базе знаний

Для получения итогового диагноза были реализованы различные модели представления знаний: таблицы референтных значений и экспертных оценок и продукционная модель.

Рассмотрим пример реализацию этих знаний в разработанной с использованием СУБД PostgreSQL базы знаний.

Таблица референтных значений хранит информацию с граничными значениями «норма-патология». Она используется для оценки результатов анализов крови и метода “ШОКС”, который реализован с помощью функций и триггеров.

Экспертная оценка также представлена в виде таблицы. В ней находятся значение важности определенного параметра, полученного на основании мнения эксперта. Этот метод используется для оценки клинического состояния и ЭКГ.

Продукционные модели представляют знания в виде структуры: если (условие), то (действие). Условия и действия для этого метода хранятся в таблице. Сам же метод реализован с помощью функций и триггеров. В отличие от референтных значений этот метод позволяет встроить в решающее правило любое действие (например, нормализовать значения или сгенерировать обобщенную оценку для фиксированного уровня диагностики)

Для работы с базой знаний был разработан интерфейс, который позволяет выполнять следующие операции: вносить, изменять, добавлять, отображать и удалять компоненты базы знаний (см. рисунок 2).

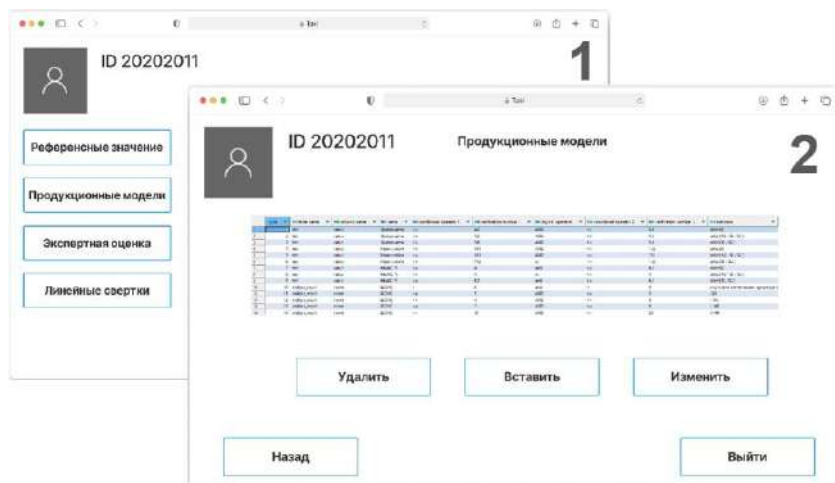


Рис.2 - Пример спроектированного интерфейса
(1 – страница выбора реализованных моделей представления знаний,
2- страница работы с производственной моделью)

Важным этапом работы является проверка корректной работы БЗ. Эта проверка выполнялась на реальных обезличенных данных пациентов. Пример расчета диагностических индексов с использованием разработанной БЗ приведен на рисунке 3.

123 Пациент	123 Уровень	123 Вес	123 Вывод
2 Гормональный анализ крови	2	1,129	1 Общих анализов гормонов не в норме
2 Общий анализ крови	2	1,059	2 Общих анализов крови не в норме
2 ЭКГ	2	0,56	Обнаружено 4 нарушения ЭКГ
2 Клиническое состояние	2	0,1967	Обнаружено 3 нарушения
2 Функциональная оценка	1	0,56	Есть проблем с Функциональная оценка
2 Лабораторные тесты	1	1,094	Есть проблем с Лабораторные тесты
2 Осмотр	1	0,1967	Нет проблем с осмотром
2 ШОКС	1	0,75	III ФК
2 Диагноз	0	0,6502	Выявлена ХСН

Рис.3 - Результат работы базы знаний

Заключение

Разработанный сегмент базы медицинских знаний стал важным шагом к систематизированному хранению и обработке информации, полученной в ходе разработки информационной системы «Регистр хронической сердечной недостаточности». Дальнейшая разработка базы знаний будет заключаться в увеличении методов представления знаний и

ее наполнения, а также проверки корректности работы на увеличенном массиве пациентов.

Литература:

1. Hak F, Guimarães T, Santos M. (2022) Towards effective clinical decision support systems: A systematic review // PLoS ONE. – 2022. – Vol. 17(8). – doi:10.1371/journal.pone.0272846.
2. Sadegh-Zadeh, K. (2011). Types of Medical Knowledge. Handbook of Analytic Philosophy of Medicine. – P. 443–458. – doi:10.1007/978-94-007-2260-6_10.
3. Riaño D, Peleg M, Ten Teije A. Ten years of knowledge representation for health care (2009-2018): Topics, trends, and challenges. Artif Intell Med. – 2019. – Vol.100. – doi:10.1016/j.artmed.2019.10171.
4. PostgreSQL: About [Электронный ресурс]. – URL: <https://www.postgresql.org/about/> (дата обращения: 20.05.2024).

РАЗРАБОТКА ДИАГНОСТИЧЕСКОГО КАЛЬКУЛЯТОРА ДЛЯ ОЦЕНКИ СТАДИИ КОЛОРЕКТАЛЬНОГО РАКА

Д.К. Кочнев, И.В. Осипенко, И.Н. Яковина

**Новосибирский государственный технический университет,
г. Новосибирск, d.kochnev.2019@stud.nstu.ru**

Научный руководитель: Яковина И.Н., к.т.н.

В работе представлен разработанный диагностический калькулятор для оценки стадии колоректального рака, проведённый анализ аналогов и требования к приложению. Разработанное веб-приложение позволяет работать с данными, вводимыми пользователем.

The paper presents the result of developing a diagnostic calculator for assessing the stage of colorectal cancer. An analysis of analogs was conducted, and requirements for the application were formulated. The developed web application allows working with data entered by the user.

Для диагностики заболеваний требуется анализировать большой объем данных. При определении стадии болезни одинаковый набор лабораторных показателей часто по-разному трактуется врачом, и восстановить его рассуждения для создания моделей поддержки доврачебных решений бывает затруднительно. Современные технологии облегчают сбор экспертных данных и позволяют создавать модели первичной диагностики с помощью методов машинного обучения. Обученные модели с высокой точностью способны

классифицировать наборы показателей, что и составляет процесс первичной диагностики заболеваний.

Диагностические калькуляторы, использующие различные клинические и лабораторные показатели, помогают определять тяжесть заболевания, оценивать вероятность осложнений и уточнять долгосрочный прогноз. Это позволяет врачам разрабатывать оптимальные планы лечения, выбирать необходимые лекарства и, при необходимости, планировать хирургическое вмешательство в тяжелых случаях. Примеры успешных диагностических калькуляторов, такие как AJCC Colorectal Cancer Staging Calculator [1], показывают важность этих инструментов в медицине.

Разрабатываемый калькулятор также предназначен для оценки стадии колоректального рака и предоставляет врачам более точный инструмент для диагностики и прогнозирования. В рамках данной работы стояла задача разработки диагностического калькулятора для оценки стадии колоректального рака.

При выборе языка программирования и библиотеки для создания приложения учитывались несколько ключевых критериев. Во-первых, язык должен быть высокопроизводительным и справляться с большими объемами данных. Во-вторых, библиотека должна предоставлять обширные инструменты для работы с машинным обучением и нейронными сетями. В-третьих, она должна поддерживать модели, созданные в Orange 3.

Python и библиотека Streamlit удовлетворяют этим требованиям и были выбраны для создания данного веб-приложения. Streamlit — это открытая библиотека Python для разработки интерактивных веб-приложений, предоставляющая инструменты для быстрой разработки интерфейсов, визуализации данных и работы с моделями машинного обучения и нейронными сетями [2, 3].

Отметим, что Streamlit не является полноценным бэкендом; его основная задача — создание интерактивного интерфейса и визуализация данных. Однако он легко интегрируется с другими библиотеками Python для работы с данными и машинным обучением. Использование Streamlit оправдано для данного приложения, так как оно не требует сложной бэкенд-логики и сосредотачивается на взаимодействии с моделями машинного обучения. Возможности Streamlit позволяют создать простое, понятное приложение, поддерживающее работу с технологиями машинного обучения. Библиотека предоставляет широкий спектр элементов для ввода и вывода данных, а также функции для работы с текстовой информацией,

отображения групп данных, отображения графиков и диаграмм, а также интерактивные элементы ввода данных.

В начальной фазе исследования были проанализированы аналогичные решения, широко применяемые в медицинском сообществе. На основе этого анализа, а также учета доступных обученных моделей машинного обучения, были сформулированы требования к инструментам разработки, включая выбор типа приложения, языка программирования и библиотек.

Веб-приложение разработано с помощью языка программирования Python и библиотеки Streamlit. Были изучены основные компоненты библиотеки Streamlit, а также ее основные концепции и принципы функционирования. Кроме того, были рассмотрены более сложные возможности, предоставляемые данной библиотекой. На основе анализа аналогичных разработок и диагностических моделей было разработано веб-приложение (Рисунок 1).

Разработанное веб-приложение предоставляет пользователям возможность вводить данные о пациентах, получать результаты работы 5 различных моделей машинного обучения и просматривать дополнительную информацию о работе этих моделей.

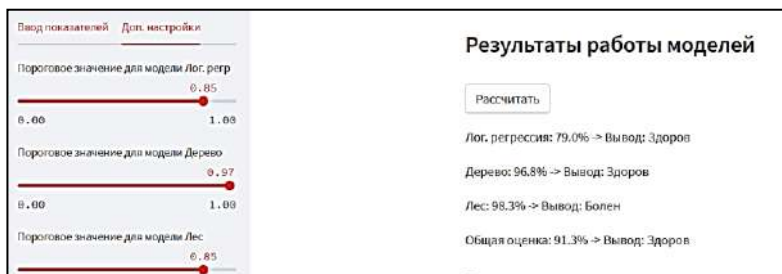


Рис.1 – Пример работы веб-приложения.

Долгосрочные перспективы развития проекта включают дальнейшее расширение функционала и интеграцию с внешними базами данных.

Литература:

1. AJCC Cancer Staging Manual [Электронный ресурс]. – URL: <https://www.facs.org/quality-programs/cancer-programs/american-joint-com-mittee-on-cancer> (дата обращения 08.11.2024).
2. Streamlit - API Reference [Электронный ресурс]. – URL: <https://streamlit-docs.netlify.app/library/api-reference> (дата обращения: 08.11.2024).

3. Python [Электронный ресурс]. – URL: <https://www.python.org/about/> (дата обращения: 08.11.2024).

АВТОМАТИЗИРОВАННАЯ СИСТЕМА ОЦЕНКИ ПСИХОФИЗИОЛОГИЧЕСКОЙ РЕАКЦИИ ЧЕЛОВЕКА НА СВЕТ

Р.А. Жакупов

**Томский политехнический университет, ИШНПТ, 4В41,
г. Томск, raz12@tpu.ru**

**Научный руководитель: Туранов С. Б., канд. техн. наук, доцент
ОМ ИШНПТ ТПУ**

В этой статье рассматривается разработка автоматизированной системы для оценки психофизиологических реакций человека на световые воздействия, основанной на методах компьютерного зрения. Описан алгоритм, анализирующий частоту морганий, диаметр зрачков и коэффициент открытости глаз для выявления признаков усталости и стресса. Система может применяться для адаптивного освещения, чтобы способствовать повышению продуктивности и улучшению самочувствия.

This article discusses the development of an automated system for assessing human psychophysiological responses to light exposure based on computer vision methods. It describes an algorithm that analyzes blink frequency, pupil diameter, and eye aspect ratio (EAR) to identify signs of fatigue and stress. The system can be applied to adaptive lighting to promote productivity and improve well-being.

В последние годы наблюдается значительный рост использования технологий компьютерного зрения (КЗ). Оно находит применение в разных областях, начиная от медицины и заканчивая космическими спутниками [3]. С 2024 по 2029 года среднегодовой темп роста мирового рынка КЗ составит 17,72% [2]. В России также уделяется особое внимание развитию КЗ в рамках национальной программы «Цифровая экономика РФ» и Национальной стратегии развития [4].

Одним из актуальных направлений для внедрения КЗ является управление осветительными системами. Неправильно организованное освещение способно снизить продуктивность и даже вызвать заболевания нервной и сердечно-сосудистой систем [1]. Для решения этой задачи необходимо разрабатывать алгоритмы, способные отслеживать состояние человека в режиме реального времени. Несмотря на активный интерес к этой теме, применение КЗ для адаптивного освещения остается малоизученным.

Для реализации автоматизированной системы оценки психофизиологической реакции на свет был разработан алгоритм на

языке Python с использованием библиотек OpenCV, Dlib, Pandas, Matplotlib, NumPy, SciPy и Scikit-Learn. Алгоритм определяет коэффициент открытости глаз (EAR), частоту морганий и диаметр зрачка. Эти показатели позволяют оценить уровень усталости и стрессового состояния человека.

С помощью библиотека Dlib, определяются координаты ключевых точек глаз. На основе расстояний между этими точками рассчитывается коэффициент EAR по формуле:

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|} \quad (1)$$

Посредством коэффициента EAR рассчитывается количество и частота морганий. Данные для последующего анализа записываются в CSV-файл и визуализируются в виде графиков и таблицы Excel (рисунок 1-2, таблица 1).

```

15  def eye_aspect_ratio(eye):
16      A = dist.euclidean(eye[1], eye[5])
17      B = dist.euclidean(eye[2], eye[4])
18      C = dist.euclidean(eye[0], eye[3])
19      ear = (A + B) / (2.0 * C)
20      return ear
21
22
23  def pupil_diameter(eye):
24      # Рассчитываем примерный диаметр зрачка (расстояние между самыми дальними точками глаза)
25      diameter = dist.euclidean(eye[0], eye[5])
26      return diameter

```

Рис.1. – Фрагмент кода из алгоритма с изображением функций, отвечающих за определение EAR и диаметра зрачка

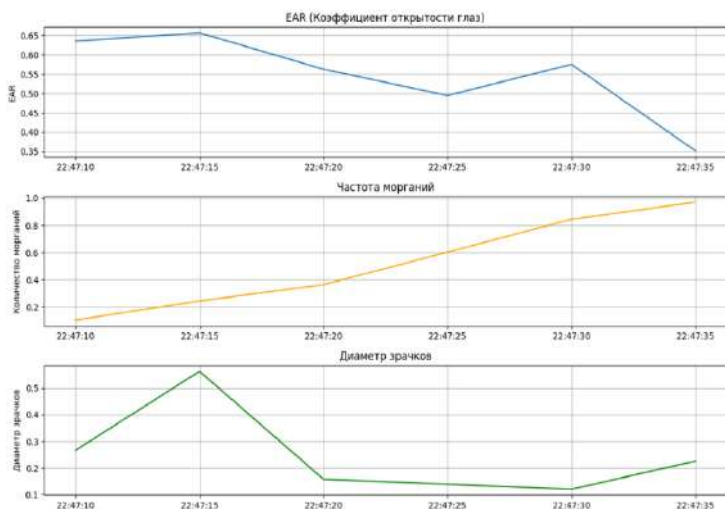


Рис.2. – Результат выполнения алгоритма в графическом виде

Таблица 1 – пример вывода данных в таблицу Excel

timestamp	EAR	blinks	pupil_diameter	state
22:47:10	0,635461	0,10495	0,266136	0
22:47:15	0,656767	0,245091	0,563801	0
22:47:20	0,562973	0,364571	0,157111	0
22:47:25	0,494689	0,603333	0,138461	0,322917
22:47:30	0,575565	0,84419	0,119331	0,247619
22:47:35	0,351651	0,970256	0,225625	0,692308

Разработанный алгоритм можно интегрировать в автоматизированные системы адаптивного освещения, которые регулируют параметры света в зависимости от состояния пользователя. Например, при снижении EAR и повышении частоты морганий система интерпретирует это, как признаки усталости и изменит яркость света или его цветовую температуру, чтобы снизить нагрузку на глаза человека.

Данный алгоритм демонстрирует возможность использования КЗ для оценки психофизиологического состояния человека с целью адаптивного управления освещением и обладает потенциалом для внедрения в интеллектуальные осветительные системы, способные улучшить условия труда и повысить уровень комфорта.

Литература:

1. Освещение и психология: влияние света на наш эмоциональный фон [Электронный ресурс] // Все о строительстве и ремонте: сайт. – 2023. – URL: <https://stroyvec.ru/osveshhenie-i-psihologiya-vliyanie-sveta-na-nash-emocionalnyj-fon/> (дата обращения: 08.11.2024).
2. Анализ размера и доли рынка компьютерного зрения – тенденция роста и прогнозы (2024-2029гг.) [Электронный ресурс] // Mordor Intelligence: сайт. – 2024. – URL: <https://www.mordorintelligence.com/ru/industry-reports/computer-vision-market> (дата обращения: 08.11.2024).
3. David A. Forsyth, Jean Ponce Computer Vision. A modern approach. // by Pearson Education. – 2003. – P.1.
4. Национальная стратегия развития искусственного интеллекта в РФ [Электронный ресурс] // Указ Президента РФ о развитии искусственного интеллекта в Российской Федерации. – 2024. – URL: https://ai.gov.ru/knowledgebase/dokumenty-po-razvitiyu-ii-v-rf/nacionalnaya_strategiya_razvitiya_iskusstvennogo_intellekta_na_period_do_2030_goda/ (дата обращения: 08.11.2024).

РАЗРАБОТКА МОДЕЛИ ПРОГНОЗА АНОМАЛИЙ В РАБОТЕ ПРОМЫШЛЕННОГО ТРАНСФОРМАТОРА

Г.А. Каталицкий

**Новосибирский государственный технический университет,
г. Новосибирск, george.ka04@list.ru**

Научный руководитель: Яковина И.Н., к.т.н.

В статье рассматриваются результаты применения методов машинного обучения для прогноза аномалий в работе промышленного трансформатора функционирующего в двух режимах, описаны полученные модели

The article discusses the results of applying machine learning methods for anomaly detection in the operation of an industrial transformer operating in two modes, and describes the models obtained.

Аварии и повреждения промышленных трансформаторов могут привести к значительным финансовым затратам эксплуатирующей их организации и экологическому ущербу. Чтобы предотвратить эти проблемы, необходима своевременная автоматическая диагностика, традиционно выполняемая специалистами на месте. Однако такой подход может быть рискованным из-за больших интервалов между

проверками. Современные системы мониторинга позволяют отслеживать параметры трансформатора и масла в режиме реального времени. Применение методов машинного обучения позволяет на основе уже собранных данных разработать модели, которые можно будет использовать для прогноза аварий. Решением данной задачи активно занимаются специалисты различных регионов, так в работе [1, 2] авторы обучали модели для предсказания общего технического состояния трансформатора.

Целью данной работы является создание моделей, способных прогнозировать уровень горючих газов в трансформаторном масле. При построении моделей предполагается, что трансформатор может работать в двух режимах:

- 1) Стабильный режим [stable operation] (SO);
- 2) Аварийный режим [accidental regime] (AR).

Входными характеристиками для разработок моделей являются:

- Температура наружного воздуха (Т), °С;
- Влажность наружного воздуха (W), %;
- Давление (P), мм.рт.ст;
- День недели(DW);
- Время суток(H), час;
- Содержание воды(H₂O), мг/л;
- Содержание водорода(H₂), мг/л;
- Ток частичных разрядов (IA, IB, IC), А;
- Мощность частичных разрядов (PA, PB, PC) мВт;

На выходе будет спрогнозировано содержание горючих газов (GG) в трансформаторном масле, мг/л.

Для того, чтобы проиллюстрировать разницу между этими режимами, рассмотрим график абсолютной концентрации GG реального промышленного трансформатора, за которым в течении семи месяцев с декабря 2023 по июль 2024 года велось наблюдение и был зафиксирован процесс дегградации масла и перехода из стабильного режима работы в аварийный. Красной линией на рисунке 1 обозначен момент перехода в аварийный режим.

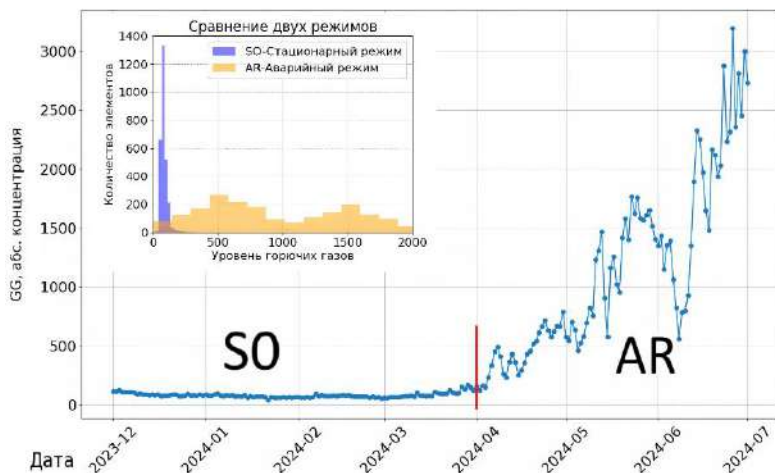


Рис. 1 – График концентрации горючих газов при стационарном и аварийном режиме.

С учетом динамики процесса электрохимических реакций прогнозирование планируется осуществлять на следующий час и на следующий день. Для этого, после проведения корреляционного анализа (см. Таблица 1), были разработаны 4 модели, оценки которых приведены в таблице 2:

- 1) Трансформатор работает в стационарном режиме, прогноз делается на следующий час (M-so1).
- 2) Трансформатор работает в стационарном режиме, прогноз делается на следующие 24 часа (M-so24).
- 3) Трансформатор работает в аварийном режиме, прогноз делается на следующий час (M-ar1).
- 4) Трансформатор работает в аварийном режиме, прогноз делается на следующие 24 часа (M-ar24).

Таблица 1 – Корреляции

<i>Пары параметров</i>	<i>Режимы</i>		<i>Примечание</i>
	<i>SO</i>	<i>AR</i>	
GG-PA,PB,P C	0.03	0.13	Среднее значение корреляции уровня горючих газов от мощности фаз.
GG-IA,IB,IC	0.04	0.08	Среднее значение корреляции уровня горючих газов от силы тока фаз. Не наблюдается статистически значимых различий между корреляциями.
GG-H2O	*	0.74	*-При работе в стационарном режиме уровень содержания воды постоянен и равен 0.00. При аварийном режиме наблюдается значительная корреляция.
GG-T	0.20	0.62	Корреляция между уровнем горючих газов и температурой в аварийном режиме обусловлена сезонными изменениям
GG-H2	0.65	0.86	Корреляция между содержанием водорода и уровнем горючих газов вызвана, тем что водород также учитывается, как горючий газ

Для прогнозирования результатов используется модель многослойного персептрона с входным слоем, который принимает 13 параметров, описанных в разделе “Входные данные”, два скрытых слоя по 100 нейронов, и выходной слой, который выдаёт концентрацию GG. Выборки данных были разделены на обучающую и тестовую в отношении 80 на 20 %. Для реализации модели используются библиотека keras. Для оценки полученных моделей используются следующие показатели [3]:

1. MAE - Chtlyzz f,cjk.nyfz jib,rf (Mean Absolute Error)\$

2. RMSE - Корень из средней квадратичной ошибки (Root Mean Squared Error);
3. R2 - Коэффициент детерминации (Coefficient of Determination).

Таблица 2 – Оценки полученных моделей

Модель	MAE	RMSE	R2	Обучающая выборка	Тестовая выборка
M-so1	11.21	18.77	0.49	2281	571
M-so24	18.35	29.50	0.27	2263	566
M-ar1	193.06	434.11	0.76	1717	430
M-ar24	253.00	440.66	0.72	1699	425

На основании полученных результатов можно отметить, что коэффициент детерминации, отражающий степень схожести между фактическими и предсказанными значениями, выше при работе трансформатора в аварийном режиме. Кроме того, следует подчеркнуть, что этот коэффициент не демонстрирует значительных различий между прогнозами на следующий час и на следующие сутки в условиях аварийного режима. Для разработки более точных моделей планируется расширить набор методов и горизонт наблюдения за трансформатором.

Литература:

1. *Khalyasmaa A.I., Matrenin P.V., Eroshenko S.A.* Assessment of Power Transformer Technical State Using Explainable Artificial Intelligence // *Problemele Energeticii Regionale*. – 2024. – Vol. 4. – N. 64. – P. 9.
2. *Dhini A., Faqih A., Kusumoputro B., Surjandari I., Kusiak A.* Data-driven Fault Diagnosis of Power Transformers using Dissolved Gas Analysis (DGA). *International Journal of Technology*. – 2020. – Vol. 11(2). – P. 388–399. doi: /10.14716/ijtech.v11i2.3625.
3. *Hodson T. O.* Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not // *Geoscientific Model Development Discussions*. – 2022. – T. 2022. – C. 1-10.

АНАЛИЗ ДАННЫХ МОНИТОРИНГА СОСТОЯНИЯ ПРОМЫШЛЕННЫХ ТРАНСФОРМАТОРОВ

В.В. Еремеева, Н.Г. Бурлаков

**Новосибирский государственный технический университет,
г. Новосибирск, vikavladim@mail.ru**

**Научный руководитель: Яковина И.Н., к.т.н.,
доцент кафедры ВТ НГТУ**

В данной статье сравниваются режимы работы двух промышленных трансформаторов, находящихся в одной локации. Проводится анализ суточной и недельной динамики, влияния погодных условий и возможности описать выявленные зависимости с помощью линейной регрессионной модели.

This article compares the operating modes of two industrial transformers located in the same location. An analysis is carried out of daily and weekly dynamics, the influence of weather conditions and the possibility of describing the identified dependencies using a linear regression model.

Мониторинг трансформаторов Сибирского отделения РЖД, которые находятся в труднодоступных регионах, является одним из приоритетных направлений модернизации инфраструктуры электрических сетей организации. Их обслуживание и ремонт является очень дорогостоящим и динамическое наблюдение с использованием методов автоматической диагностики и прогнозирования существенно экономит расходы. В работе используются данные мониторинга двух промышленных трансформаторов с разными состояниями работоспособности, находящихся в одной локации.

Данные собирались с четвертого и пятого трансформатора, поэтому примем за обозначение трансформаторов следующие аббревиатуры: Т4 и Т5. Анализируемые данные обозначаются следующим образом: Н₂ – содержание водорода в масле трансформатора; Н₂О – содержание воды в масле трансформатора; ГГ – содержание горючих газов в масле трансформатора; ЧР – частичный разряд; Мах – максимальный частичный разряд за текущее время среди фаз А, В, С; Т – температура; U – влажность; Р – давление.

Для анализа были предоставлены данные за период с мая по август 2020 г. включительно. По некоторым значениям данные собирались только с середины июля, в следствие чего, в качестве исследуемого промежутка времени были взяты результаты мониторинга с 16 июля 2020 по 03 сентября 2020 (период непрерывного сбора всех параметров). Это летний режим работы со стабильной температурой и влажностью.

В ходе анализа набора данных с целью выявления суточных и недельных паттернов для двух трансформаторов были получены

графики средних значений H_{20} , H_2 , ГГ, Мах с явно выраженным характером циклических изменений в течении периода наблюдений. Результат представлен на Рис. 1.

По данным графиков можно сделать вывод о влиянии времени суток и дня недели на некоторые параметры и выделить те параметры, которые зависят от времени суток и дня недели в целом или на фиксированном интервале периода. К первой группе можно отнести суточное H_{20} , суточное и недельное H_2 и ГГ с небольшой абберацией на суточном и суточное Мах, так как колебания для этих значений имеют одинаковую тенденцию для обоих трансформаторов. Недельные циклы H_{20} и Мах имеют не явно выраженную цикличность, которая слабо зависит от дня недели, по обоим графикам наблюдается сдвиг на половину фазы между трансформаторами.

Суточный график Мах отражает разницу в техническом состоянии трансформаторов. Динамика значения параметра Т4 имеет явно выраженный нестабильный характер изменения величины по сравнению с Т5. Это может свидетельствовать о нештатном режиме работы 4Т и иллюстрировать частые скачки частичных разрядов по трем фазам.

Для выявления взаимосвязи между исследуемыми параметрами были получены матрицы коэффициентов корреляции. В Таблице 1 представлены характерные пары метрик по двум трансформаторам. По приведенным данным видна разница, которая очевидно обусловлена нештатным режимом работы Т4.

Таблица 1 – Корреляционный анализ

Пары параметров	H_2-H_{20}	Мах–ГГ	Мах–U	Мах–Т	Мах–Р
Т4	0.04	0.19	0.12	-0.15	-0.21
Т5	-0.66	-0.31	-0.63	0.60	0.38

После проведения корреляционного анализа были построены линейные регрессионные модели с использованием библиотеки `sklearn.linear_model`. По данным моделирования, приведенным в Таблице 2, явно прослеживается разница в метрике R^2 для Т4 и Т5, т.е. возможно с довольно высокой точностью предсказывать значения только для пятого трансформатора. Это, вероятно, связано с техническим состоянием данных трансформаторов. И можно сделать вывод, что Т5 находится в более работоспособном состоянии и его поведение более системно и предсказуемо, тогда как поведение Т4 предсказать невозможно, что свидетельствует о его нештатном режиме работы [2].

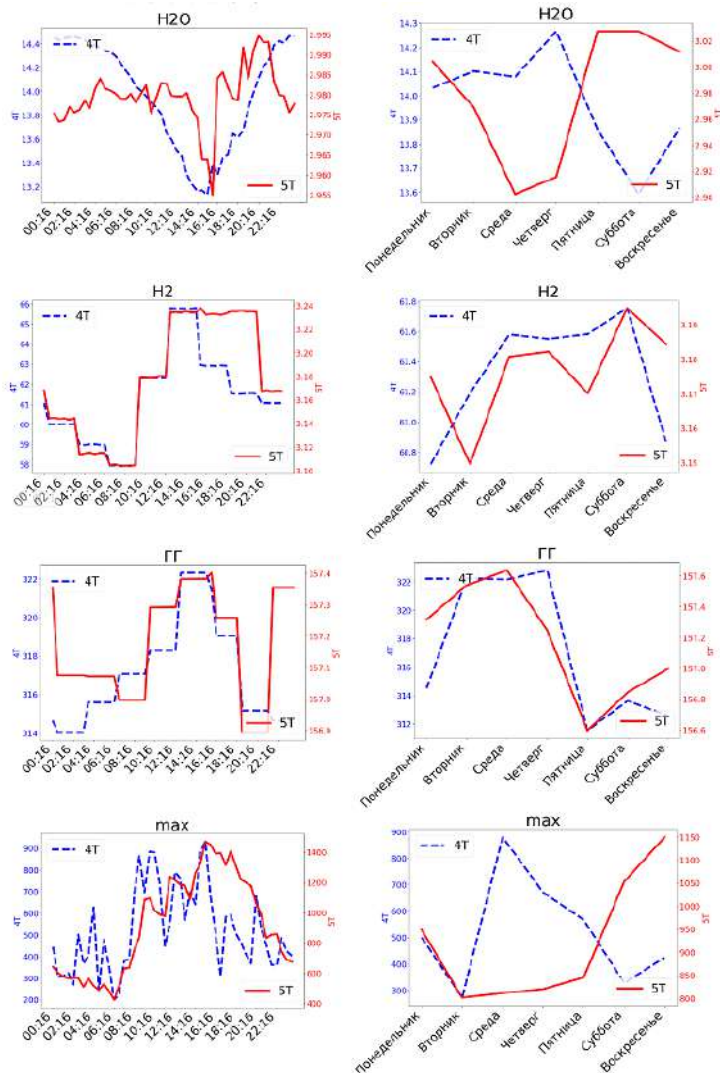


Рис. 1 – Суточная и недельная динамика по метрикам

Примечание: левая ось – 4Т, правая ось – 5Т

Таблица 2 – Параметры и оценки моделей

Модели		41	42	51	52
Метрики	MAE	0.43	0.46	0.48	0.48
	RMSE	0.97	0.97	0.67	0.67
	R2	0.07	0.07	0.47	0.47
Коэффициенты регрессионных моделей	T	-0,108	-0,109	0,280	0,283
	U	-0,097	-0,092	-0,341	-0,335
	P	-0,138	-0,138	0,159	0,153
	H2	-0,039	-0,039	-0,123	-0,135
	H2O	-0,146	-0,144	-0,142	-0,148
	ГГ	0,181	0,180	0,016	0,020
	hour	-	0,009	-	-0,002
	day_of_week	-	0,012	-	0,056
	Добавочный	-0.001	-0.001	-0.004	-0.003

Вывод

В ходе анализа данных мониторинга состояния промышленных трансформаторов были выявлены суточные и недельные закономерности изменений характеристик трансформатора. Были составлены многомерные линейные регрессионные модели. Подтверждена гипотеза нештатного режима работы трансформатора Т4 и определено влияние, которое этот режим оказывает на характеристики и их циклы изменения.

Литература:

1. *Lyutikova M.N., Korobeynikov S.M., Ridel A.V.* Method for Reducing Sediment Formation in Transformer Oil // IEEE Transactions on Dielectrics and Electrical Insulation. – 2024. – Vol. 31, N 4. – P. 2208-2215.
2. *Khalyasmaa A.I., Matrenin P.V., Eroshenko S.A.* Assessment of Power Transformer Technical State Using Explainable Artificial Intelligence // Problemele Energeticii Regionale. –2024. – Vol. 4, N 64. – P. 9.
3. *Korobeynikov S.M., Ridel A.V., Ovsyannikov A.G. et al.* Streamers in transformer oil initiated by partial discharge in a bubble at low AC electric fields // Electr Eng. – 2024. – Vol. 106. – P. 5219–5228.

РАЗРАБОТКА МЕТОДА ОЦЕНКИ РИСКА ВОЗНИКНОВЕНИЯ ДТП С УЧАСТИЕМ ПЕШЕХОДОВ НА ОСНОВЕ ДИНАМИЧЕСКОЙ ИДЕНТИФИКАЦИИ СОСТОЯНИЯ СРЕДЫ

А.С. Гребенкин

Новосибирский государственный технический университет,

г. Новосибирск, a.grebenkin.2019@stud.nstu.ru

Научный руководитель: Ганелина Н.Д., к.т.н.

В работе представлен метод оценки риска возникновения дорожно-транспортных происшествий с участием пешеходов, основанный на сопоставлении текущих факторов внешней среды с историческими данными. Предложен алгоритм расчета меры близости погодных и световых условий, учитывающий численные и категориальные показатели. С помощью методов статистического анализа определено пороговое значение меры близости для выделения зон повышенного риска.

The algorithm for traffic accidents risk assessing based on comparative analysis of the current environmental factors with historical data on accidents with pedestrians is presented in the paper. The method for the proximity measure of weather and lighting conditions, taking into account numerical and categorical indicators, calculating is suggested. The threshold value of the proximity measure for identifying high-risk zones is determined using statistical analysis methods.

Проблема дорожно-транспортных происшествий с участием пешеходов является значимой для обеспечения безопасности дорожного движения. В России доля таких ДТП составляет более 30% от общего числа происшествий [1]. Для снижения их количества важно не только анализировать статистические данные, но и определять зоны повышенного риска на основе текущих условий.

Существующие исследования подтверждают значительное влияние погодных условий на вероятность возникновения ДТП. Так, в работе [2] показана корреляция между среднесуточным уровнем осадков и числом происшествий. Также установлено [3], что для обеспечения безопасности пешеходов необходимо учитывать не только текущие условия, но и специфику места происшествия. Исследования с применением байесовских сетей [4] позволили выявить взаимосвязь между различными факторами риска и вероятностью ДТП с участием пешеходов.

На основе анализа существующих исследований [2-4] были выделены факторы, информация по которым доступна как в исторических данных о ДТП, так и в текущий момент времени через внешние сервисы. К таким факторам относятся погодные условия (температура, осадки, видимость, скорость ветра, облачность) и

естественная освещенность (время суток относительно восхода/заката солнца).

Предлагаемый метод основан на сопоставлении текущих условий с условиями, при которых происходили ДТП ранее. Для численного сравнения условий разработан способ расчета меры близости, учитывающий три компонента.

Первый компонент отвечает за сравнение численных показателей погоды и использует евклидово расстояние между векторами стандартизованных значений (x и y - векторы показателей текущих и исторических условий, включающие температуру воздуха, уровень осадков, видимость, скорость ветра и облачность):

$$D(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (1)$$

Стандартизация каждого показателя производится по формуле (μ - среднее значение показателя, σ - стандартное отклонение):

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

Это позволяет привести все показатели к единому масштабу и учесть их естественную вариативность.

Второй компонент оценивает близость времени суток с учетом естественного освещения. Для этого разработан следующий алгоритм:

1. Вычисляется длительность дня $T_d = T_z - T_v$ и ночи $T_n = T_v - T_z$, где T_z - время заката, T_v - время восхода.
2. Определяются временные интервалы между текущим временем T и временем восхода $T_{tv} = T - T_v$, а также между текущим временем и закатом $T_{tz} = T - T_z$.
3. Производится стандартизация интервалов.

Для дневного времени:

$$T_{ст_tv} = \frac{2 * T_{tv}}{T_d} \quad (3)$$

$$T_{ст_tz} = \frac{2 * T_{tz}}{T_d} \quad (4)$$

Для ночного времени:

$$T_{\text{СТ_ТВ}} = \frac{2 * T_{\text{ТВ}}}{T_{\text{Н}}} \quad (5)$$

$$T_{\text{СТ_ТЗ}} = \frac{2 * T_{\text{ТЗ}}}{T_{\text{Н}}} \quad (6)$$

4. В качестве меры близости времени выбирается наименьшее из значений $|T_{\text{СТ_ТВ}}|$ и $|T_{\text{СТ_ТЗ}}|$.

Третий компонент учитывает категориальные значения погодных условий (ясно, пасмурно, дождь, туман, снегопад, метель). Для их сравнения используется бинарный показатель: 1 - если условия совпадают, 0 - если различаются. Данный компонент важен, так как некоторые погодные явления, например туман, могут значительно повышать риск ДТП независимо от численных показателей видимости и влажности.

Итоговая мера близости вычисляется по формуле (w_1 , w_2 , w_3 - весовые коэффициенты компонентов, подобранные эмпирическим путем для обеспечения сбалансированного учета всех факторов, D - евклидово расстояние стандартизованных численных показателей, B - бинарный показатель совпадения категориальных значений):

$$M = w_1 * D + w_2 * \min(T_{\text{СТ_ТВ}}, T_{\text{СТ_ТЗ}}) + w_3 * B \quad (7)$$

Полученная мера близости формируется как взвешенная сумма трех компонентов с эмпирически подобранными весовыми коэффициентами. Для определения порогового значения был проведен статистический анализ распределения данной величины на исторических данных за последний год. По результатам было определено, что мера близости принимает значения от 0.00025 до 21.6, а минимальные значения в каждой группе сравнений находятся в диапазоне от 0.00025 до 0.0066.

На основе полученных данных в качестве порогового было выбрано значение 0.025, что обеспечивает гарантированное наличие зон риска при любых условиях и выявление в среднем 25 зон риска при каждом запросе. Тестирование на статистических данных Новосибирского района показало стабильную работу метода при различных погодных и световых условиях.

Предложенный алгоритм реализован в мобильном приложении для операционной системы Android. Функционал приложения позволяет просматривать статистику ДТП на карте, применять различные фильтры для анализа данных и получать информацию о зонах риска в текущих условиях. На рисунке 1 представлен интерфейс приложения с

отображением зон риска ДТП, сформированных на основе разработанного метода.

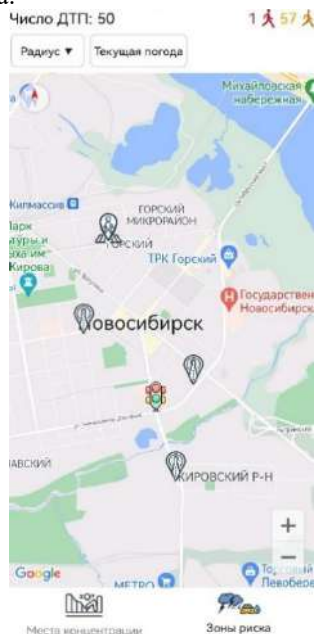


Рис. 1 – Интерфейс мобильного приложения с отображением зон риска ДТП

Разработанный метод может быть использован в навигационных системах для предупреждения водителей об опасных участках в конкретных условиях движения. В дальнейшем планируется расширение набора учитываемых факторов и совершенствование алгоритма оценки их влияния на риск возникновения ДТП.

Литература:

1. МВД России. Наезды на пешеходов составляют треть всех дорожно-транспортных происшествий в России [Электронный ресурс] // МВД РФ. – URL: <https://мвд.рф/news/item/21544340> (дата обращения: 07.11.2024)
2. Yannis G., Karlaftis M.G. Weather effects on daily traffic accidents and fatalities: a time series count data approach // Proceedings of the 89th Annual Meeting of the Transportation Research Board. – 2010. – С. 10-14.
3. Hussain Q. et al. The relationship between impact speed and the probability of pedestrian fatality during a vehicle-pedestrian crash: A

systematic review and meta-analysis // Accident Analysis & Prevention. – 2019. – Т. 129. – С. 241–249.

4. Jianjun Y. et al. Analysis of the causes of pedestrian-vehicle traffic accidents based on Bayesian networks. / Y. Jianjun, L. Yu, S. Hongbo, C. Yimeng // Preprint.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ОПТИМИЗАЦИИ ПО СТОИМОСТНОЙ ФУНКЦИИ НА РАСПОЛОЖЕНИЕ КОРНЕЙ СИСТЕМЫ, СОСТОЯЩЕЙ ИЗ ПИД-РЕГУЛЯТОРА И ОБЪЕКТА ВТОРОГО ПОРЯДКА НА ГРАНИЦЕ УСТОЙЧИВОСТИ

М.В. Скворцов

**Новосибирский государственный технический университет,
г. Новосибирск, skvmat03@gmail.com**

Научный руководитель: Саблина Г.В., к.т.н., доцент

В данной работе исследовано влияние различных стоимостных функций на оптимизацию параметров ПИД-регулятора, исследовано влияние стоимостной функции оптимизации на корневой портрет системы. Исследование проведено относительно корней объекта, расположенных на границе устойчивости.

In this paper, the influence of various cost functions on the optimization of the parameters of the PID controller is investigated, the influence of the cost optimization function on the root portrait of the system is investigated. The study was conducted with respect to the roots of the object on the boundary of stability.

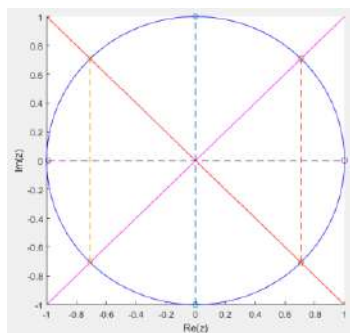


Рис.1 – Изображение наборов корней на z-плоскости

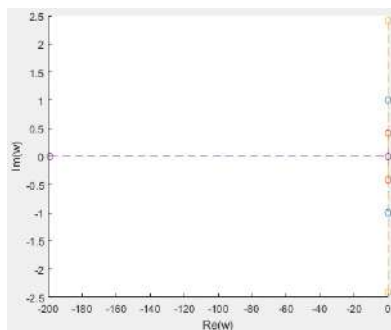


Рис.2 – Изображение соответствующих наборов корней на ω -плоскости

Таблица 1 – Таблица соответствия z- и ω -корней и конечные характеристические уравнения ОУ

z_1	z_2	ω_1	ω_2	$A(s)$	N
0-1j	0+1j	0-1j	0+1j	s^2+1	1
$0,5\sqrt{2}-0,5\sqrt{2}j$	$0,5\sqrt{2}+0,5\sqrt{2}j$	$0-(\sqrt{2}-1)j$	$0+(\sqrt{2}-1)j$	$s^2+0,1716$	2
$-0,5\sqrt{2}-0,5\sqrt{2}j$	$-0,5\sqrt{2}+0,5\sqrt{2}j$	$0-(\sqrt{2}+1)j$	$0+(\sqrt{2}+1)j$	$s^2+5,8284$	3
-1(-0,99)	1	$-\infty(-199)$	0	$s^2+199s+0$	4

Для оптимизации системы будут использованы следующие стоимостные функции (СФ):

$$f_{\text{Cost}}(1) = \int t \cdot |e(t)| dt; f_{\text{Cost}}(2) = \int e^2(t) dt; f_{\text{Cost}}(3) = \int t \cdot |e(t)| dt; \quad (1)$$

$$f_{\text{Cost}}(4) = \int t \cdot e^2(t) dt$$

Синтез модели объекта управления:

$$W_{\text{Obj}}(s) = \frac{q_B}{s^2 + ps + q_A} \quad (2)$$

Для наборов корней N1, N2, N3: $q_B = q_A$; для набора корней N4: $q_B = 1, q_A = 0$ (из $A_4(s) = s^2 + 199s + 0$) (в случае $q_B = q_A$: $W_{\text{Obj}}(s) = 0$).

Для аналитического поиска корней, после нахождения оптимальных параметров регулятора, необходимо привести поиск характеристического уравнения всей системы (передаточная функция объекта сведена к общему виду):

$$W'(s) = \left(\frac{K_i}{s} + K_p + \frac{K_d s}{0,01s + 1} \right) \cdot \frac{q}{s^2 + ps + q} \quad (3)$$

$$W(s) = \frac{W'(s)}{1 + W'(s)} = \frac{B(s)}{A(s)} \quad (4)$$

$$B(s) = q \cdot ((0,01 \cdot K_p + K_d) \cdot s^2 + (K_p + 0,01 \cdot K_i) \cdot s + K_i) \quad (5)$$

$$A(s) = 0,01 \cdot s^4 + (1 + 0,01 \cdot p) \cdot s^3 + (p + 0,01 \cdot q + q \cdot (0,01 \cdot K_p + K_d)) \cdot s^2 + (q \cdot (1 + K_p + 0,01 \cdot K_i)) \cdot s + K_i \cdot q \quad (6)$$

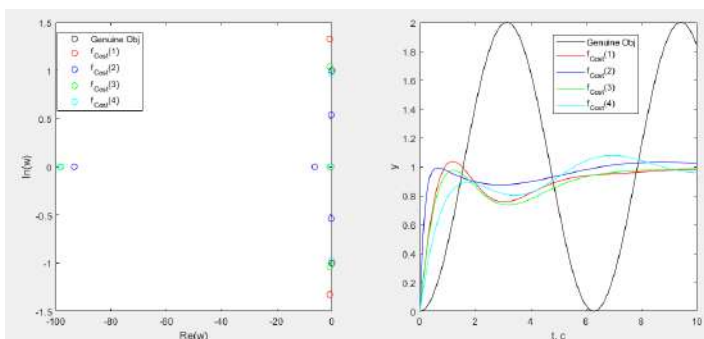


Рис.3 – Корневые портреты системы ($\omega(N1)$) и соответствующие им графики на выходе системы в зависимости от стоимостных функций

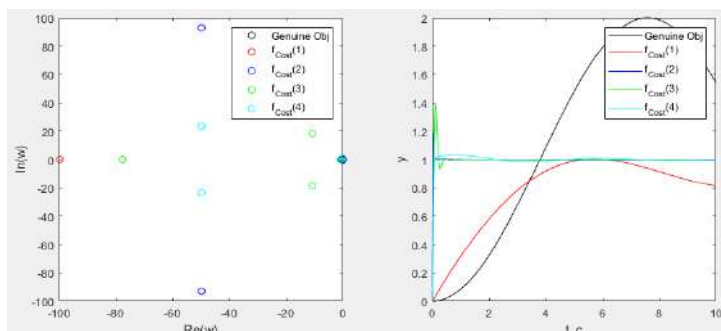


Рис.4 – Корневые портреты системы ($\omega(N2)$) и соответствующие им графики на выходе системы в зависимости от стоимостных функций

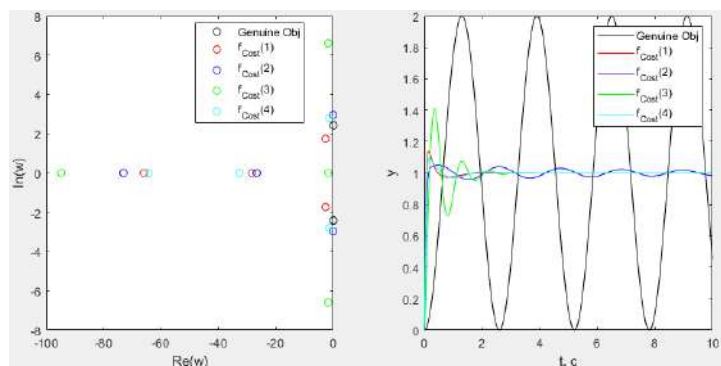


Рис.5 – Корневые портреты системы ($\omega(N3)$) и соответствующие им графики на выходе системы в зависимости от стоимостных функций

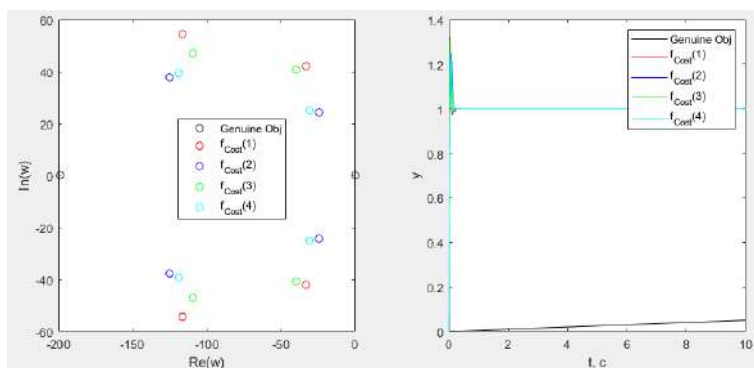


Рис.6 – Корневые портреты системы ($\omega(N4)$) и соответствующие им графики на выходе системы в зависимости от стоимостных функций

Таблица 2 – Таблица коэффициентов ПИД-регулятора в зависимости от набора корней объекта и применяемой СФ

Вид СФ	(1)	(2)	(3)	(4)
(N1) K_p	2	2	1,38197	0,381966
(N1) K_i	0,84702	2	0,75794	0,956291
(N1) K_d	2	6,23607	2	1,38197
(N2) K_p	0,381966	272,936	2161,99	83,3599
(N2) K_i	0,381966	0,445517	2043,56	282,875
(N2) K_d	2	647,218	107,301	178,354
(N3) K_p	18,3111	1,01534	7,64614	10,2551
(N3) K_i	34,6468	29,4164	13,4478	35,1744
(N3) K_d	3,94762	3,38197	0,872854	3,98985
(N4) K_p	12871,5	9258,05	13872,4	10864,2
(N4) K_i	474201	202196	462659	245070
(N4) K_d	20,1969	12,9809	12,1351	10,3582

В заключение необходимо отметить, что стоимостные функции, включающие параметр времени, задают хотя бы один быстрый корень, объясняется это тем, что в таком случае максимально быстро минимизируется ошибка, а минимизация ошибки минимизирует стоимостную функцию. А также, для колебательных объектов оптимальные корни зависят: «Больше частота колебаний – быстрее (далее) (оптимальные) корни»; для системы с вещественными корнями ОУ, были получены симметричные корневые портреты.

Литература:

1. Цифровые системы управления. Сборник задач для индивидуальных заданий: учебное пособие / Г. В. Саблина – Новосибирск: Изд-во НГТУ, 2020. – 70 с.
2. Численная оптимизация замкнутых систем автоматического управления в программе VisSim: новые структуры и методы: монография / В.А. Жмудь. – Новосибирск: Изд-во НГТУ, 2016. – 259 с.
3. Теория автоматического регулирования: Учеб. пособие / А.С. Востриков, Г.А. Французова. – Новосибирск: Изд-во НГТУ, 2006. – 368 с.

РАЗРАБОТКА И ВНЕДРЕНИЕ ТЕРМОПРОФИЛИРУЮЩЕЙ СИСТЕМЫ ОНЛАЙН МОНИТОРИНГА НА БАЗЕ ГЛУБОКОВОДНОГО НЕЙТРИННОГО ТЕЛЕСКОПА

Ю.А. Воробьева, Е.А. Уварова, И.А. Потапова
Российский государственный гидрометеорологический
университет «РГГМУ»,
г. Санкт-Петербург, Россия, makarova.yulia050801@mail.ru
Научный руководитель: Асламов И.А., к.ф.-м.н.

Вода – один из основных ресурсов нашей планеты и его состояние напрямую влияет на жизнь всех организмов, которые в ней обитают. Одним из ключевых показателей, характеризующих состояние воды в водоёмах, является температура. В данной работе предпринята попытка разработки и испытания современной системы онлайн-мониторинга температуры.

Water is one of the main resources of our planet and its condition directly affects the life of all organisms that live in it. One of the key indicators characterizing the state of water in reservoirs is temperature. In this paper, an attempt is made to develop and test a modern online temperature monitoring system.

Для измерения температуры воды на Байкале используются различные методы и приборы. В частности, наиболее распространённым методом является замер температуры с помощью термометров, которые погружаются на определённые глубины [1].

Техническая часть.

Микроконтроллер ESP32-WROVER-E [2] является оптимальным выбором для создания высокопроизводительной и надёжной системы. Это оптимальное решение, поскольку единая трехпроводная шина позволяет оптимизировать системную схему. Дatalogгер термопрофилирующей системы на микроконтроллере осуществляет

сбор данных с датчиков температуры MAX31888 в определённом интервале времени, после чего сохраняет полученные данные во встроенной памяти или на съёмном накопителе.

Для обеспечения питания термокосы была использована литиевая ячейка, извлечённая из электромобиля Nissin Leaf. Изначально ячейка имела два элемента последовательно, что давало напряжение в 8,4 В. Однако, были перекоммутированы элементы параллельно, чтобы получить напряжение 4,2 В и увеличить ёмкость ячейки вдвое. Таким образом, ёмкость повысилась до 70 А·ч.

Для реализации протокола 1-Wire и работы с термокосой, была сделана переходная плата на базе микроконтроллера PIC12F1840 [3]. Этот подход позволил расширить диапазон функций и обеспечить более эффективное управление.

На рисунке 1 представлена принципиальная схема платы PIC12F1840. Для обеспечения стабильного питания термокосы был использован контроллер, который посылает сигнал PGD для включения питания. Высокий уровень сигнала активирует DC-DC преобразователь SX1308, который запитывается от литиевых батарей с напряжением 4,2 В и повышает напряжение до 14 В. Затем используется линейный понижающий преобразователь 78L12, который стабилизирует напряжение на уровне 12 В.

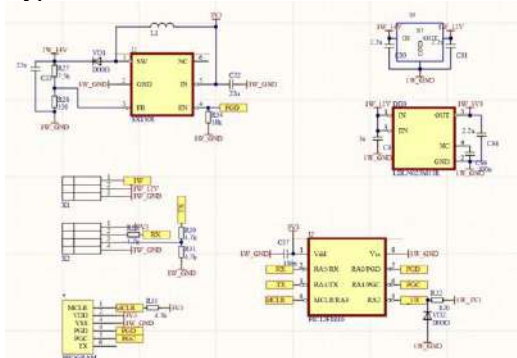


Рис. 1 – Принципиальная схема платы PIC12F1840

В ходе работы был использован дополнительный линейный стабилизатор для снижения возможных шумов, которые могут возникнуть от повышающего импульсного преобразователя SX1308.

Важно учесть, что стабильность и точность передачи данных между термокосой и контроллером являются ключевыми факторами для эффективного функционирования устройства. Для достижения этой цели, используется 12 вольт стабилизированного напряжения, которое

далее преобразуется в 3,3 вольта для линии данных 1-Wire через стабилизатор напряжения LDLN025M33R.

Контроллер термокосы имеет еще один источник питания. Контроллер также запитывается от других 3,3 вольт, которые берутся напрямую с платы ESP.

Используется трёхжильный кабель, где витая пара на экране передаёт землю и линию данных, а по экрану передается 12 В питания. Однако, длина кабеля может привести к потере качества сигнала и/или снижению стабильности обмена данными в системе. Для устранения этой проблемы в систему была введена микросхема TCA9517, которая выполняет функцию ретрансляции 1-Wire сигнала [4-5].

Для защиты датчиков от внешних воздействий, каждый датчик был не только распаян на кабель, но и помещён в латунный корпус. После этого, корпус был заполнен двухкомпонентным полиуретаном. Для калибровки датчиков использовалось оборудование – жидкостный термостат ТЕРМОТЕСТ-100-40. Благодаря этому, мы смогли добиться очень точной (до $\pm 0,01$ °C) и стабильной температуры в жидкостной среде, рисунок 2.

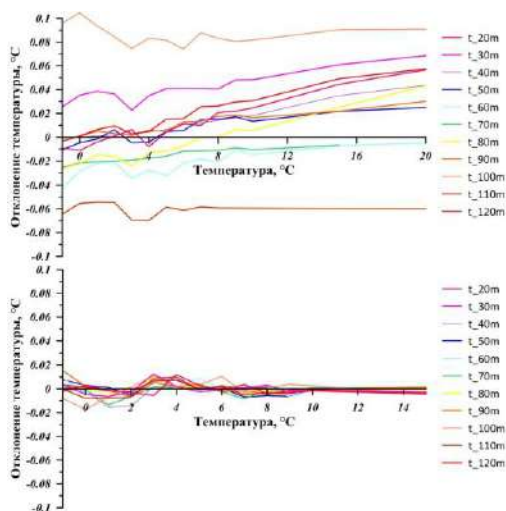


Рис. 2 – Значение температуры до и после калибровки датчиков

Экспериментальная часть.

Система термопрофилирующей системы прошла серию проверок её качества монтажа. На глубинах от 20 до 120 м, каждые 10 метров, были установлены датчики, рисунок 3, которые регистрировали температуру

в воде. После проведения измерений на датчиках были зафиксированы значения температуры на каждой глубине.

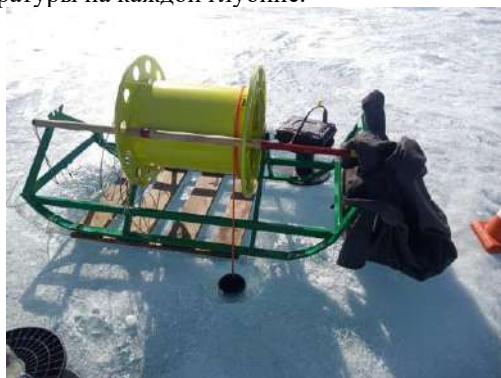


Рис. 3 – Экспериментальная установка термокосы

Согласно представленному на рисунке 4 графику, можно заметить, что работа датчиков термокосы происходит весьма стабильно без каких-либо поломок. Необходимо отметить, что такая стабильность может быть обусловлена множеством факторов и потому необходимо провести дополнительные исследования, чтобы получить полную картину работы устройства в различных условиях.

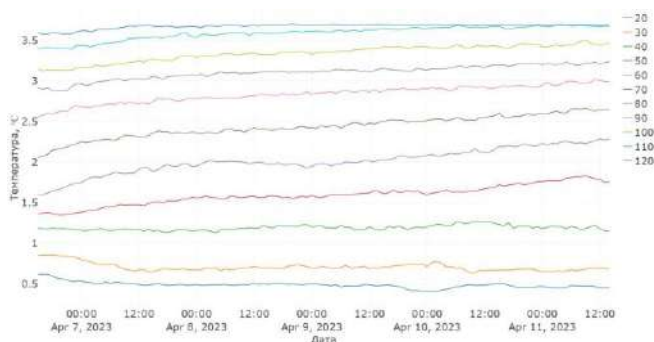


Рис. 4 – Измерение температуры за весь период исследования с усреднением по часу

Вывод.

Эта инновационная система позволяет контролировать вертикальные профили температуры воды в режиме реального времени, что ранее было затруднительно. В последующем мы планируем расширить

производство таких термокос, чтобы они могли быть использованы для изучения более широкого спектра мест на нашей планете.

Литература:

1. *Асламов И. А., Макаров М. М., Гнатовский Р. Ю., Портянская И. А.* Термопрофилирующая система онлайн-мониторинга деятельного слоя на базе глубоководного нейтринного телескопа Baikal-GVD 3ю – Иркутск: издательство Известия Иркутского государственного университета. Серия: Науки о Земле. – 2022. – Т. 42. – С. 3-15.
2. ESP32-WROVER-E & ESP32-WROVER-IE Datasheet [Электронный ресурс] / Espressif System: сайт. – URL: <https://www.espressif.com>. (дата обращения: 05.11.2022).
3. PIC12F1840 Datasheet [Электронный ресурс]/Microchip: сайт. – URL: <https://www.farnell.com/datasheets/2133777.pdf>. (дата обращения: 15.01.2023).
4. *Буталов А.В., Паклин А.И., Конюченко Н.Н. [и др.]*. Свидетельство о государственной регистрации программы для ЭВМ № 2017613439 Российская Федерация. Программа для разработки и администрирования баз данных сервера PostgreSQL "EMS SQL Management Studio for PostgreSQL": № 2016660907: заявл. 18.10.2016; опубл. 17.03.2017 / заявитель Общество с ограниченной ответственностью «ЕМС Софтваре Девелопмент». – EDN BKMLCI.
5. *Уруков С.Д.* Внедрение открытой платформы «Grafana» для аналитики time-series данных в информационную систему автоматизированных зданий / С. Д. Уруков // Молодежная школа-семинар по проблемам управления в технических системах имени А.А. Вавилова. – 2018. – Т. 1. – С. 50-53. – EDN UQHACJ.

РАЗРАБОТКА АВТОНОМНОГО ЛОГИСТИЧЕСКОГО РОБОТА

М.Д. Кабелин, Д.А. Моисеев

Новосибирский государственный технический университет,

г. Новосибирск, kabelin_nstu@bk.ru

Научный руководитель: Яковина И.Н., к.т.н.

В данной статье обсуждается задача разработки автономного логистического робота, разработанного для перемещения паллет по территории склада. Аппаратная часть робота состоит из управляющих платформ, исполнительных устройств и датчиков. Разработанная программа управления роботом протестирована, подтверждена её работоспособность.

This article discusses the task of developing an autonomous logistics robot designed to move pallets around the warehouse. The robot's hardware consists of control platforms, actuators and sensors. The developed robot control program has been tested and its operability has been confirmed.

Введение

В современном мире эффективное управление складскими процессами играет ключевую роль в успешной деятельности любого предприятия. Автоматизация складских процессов позволяет повысить производительность, сократить временные издержки и уменьшить влияние человеческого фактора, повышая эффективность логистической системы [1].

В статье представлен разработанный лабораторный стенд, моделирующий работу автономной подвижной платформы с функцией транспортировки паллет автоматизированного склада. Концепт модель робота была разработана с использованием основных особенностей решений автоматически управляемого транспорта (AGV — Automatic Guided Vehicle) [2]. Стенд предназначен для изучения различных алгоритмов управления. Базовый функционал стенда может использоваться для студентов инженерных направлений и школьников старших классов, изучающих основы робототехники.

Аппаратная часть

3Д модель разработанного логистического робота представлена на рисунке 1. Аппаратная часть робота состоит из большого количества исполнительных устройств, датчиков и микроконтроллеров, которые позволяют выполнять ряд такелажных работ. Для обеспечения корректной работы разработанный робот был оснащен комплектом различных датчиков.

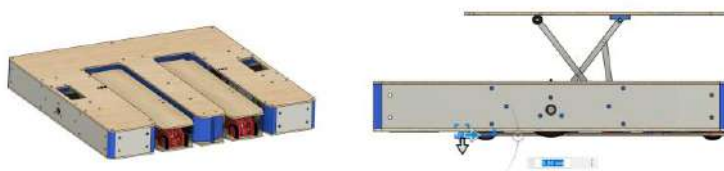


Рис. 1 – 3Д модель логистического робота

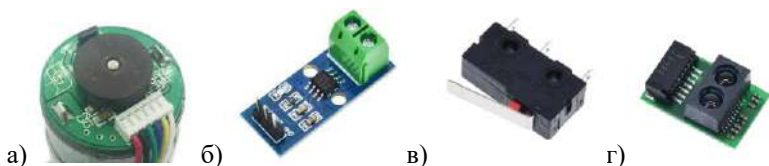
Исполнительные устройства (Рисунок 2) используются для движения шасси (мотор-редукторы постоянного тока) [3], работы подъемного механизма (моторы постоянного тока, сервоприводы).



а) мотор-редуктор «A58-555-24V-330R»; б) мотор-редуктор «5840-31ZY»; в) сервопривод «MG995»; г) микромотор «GM12-N20».

Рис. 2 – Исполнительные устройства

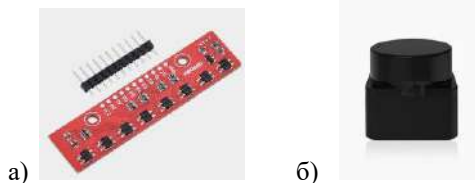
Датчики для корректной работы внутренних элементов (Рисунок 3) используются для точного позиционирования (датчик оборотов мотора), для контроля тока, протекающего через двигатели (датчики тока), определения крайних положений подъёмного механизма (концевые датчики), работы подъёмного механизма (инфракрасный датчик позволяющий определять расстояние, на которое он выдвигается).



а) датчик оборотов «A3144E TO-92»; б) датчик тока «ACS712», в) концевой датчик «KW11-3Z», г) инфракрасный датчик «GP2Y0E03»

Рис. 3 – Датчики для работы внутренних элементов

Датчики для ориентации в пространстве (Рисунок 4) используются для движения по линии и определения перекрёстков (инфракрасный датчик линии), определения положения объектов вокруг робота (ЛИДАР, который позволяет обеспечить безопасное передвижение робота и модификацию способа навигации при необходимости).



а) датчик «ECKE5000»; б) датчик «DToF LIDAR ZT-LD06»

Рис. 4 – Датчики для ориентации в пространстве

Силовой электроникой робота являются драйвера двигателей постоянного тока и реле включающие и отключающие питание всего робота. Компоненты силовой электроники изображены на рисунке 5.



а)



б)

а) драйвер «BTS7960»; б) реле электромагнитное «TRAF-D-5VDC-S-H»

Рис. 5 – Компоненты силовой электроники робота

Программно-аппаратная часть

Модуль контроллера – «кремниевый мозг» робота получает и обрабатывает данные с сенсоров и осуществляет управление драйверами актуаторов. Платы *управляющей электроники* изображены на рисунке 6. В качестве одноплатного компьютера был выбран «Raspberry Pi4 Model B» для последующего расширения функционала. В качестве микроконтроллеров управления системой восприятия и движения были выбраны два микроконтроллера «Arduino Nano». Обмен информацией между платами осуществляется по протоколу UART. Его аппаратная реализация есть на каждой плате.

Для управления подъемным механизмом, системой передвижения и фиксации вилки, шасси робота, системой слежения за контрастной линией используются двигатели постоянного тока и сервомоторы. Управление двигателем осуществляется драйвером двигателя постоянного тока со схемой Н-мост при помощи ШИМ сигнала, который интерпретируется как скважность питающего напряжения.

Система слежения за контрастной линией использует сигналы 8 аналоговых датчиков линии, установленных на 1 плате. Управление роботом осуществляется посредством задания скорости двигателям робота. Задача регулятора удерживать линию по центру датчика линии.

Вся программная часть написана на языке Arduino в среде ArduinoIDE. В начале программы проходит инициализация и настройка портов ввода/вывода, затем подаются управляющие сигналы на исполнительные устройства шасси и подъемного механизма. Мобильная платформа имеет два режима перемещения: по контрольным точкам с помощью метода конечных автоматов и по черной линии, нанесенной на напольное покрытие. Программный код будет приведен в ходе доклада на конференции.



а) Raspberry Pi4 Model B; б) Arduino Nano; в) сборка корпуса

Рис. 6 – Управляющая электроника

Заключение

В результате проделанной работы были решены следующие задачи:

1. Разработана структурная схема и компоновка лабораторного стенда, в том числе механическая и электронная части;
2. Был подобран набор аппаратных решений для логистического робота, которые позволят ему выполнять задачи перевозки грузов;
3. Написаны и отлажены алгоритмы управления мобильной платформой.

Планируется повысить степень автономности за счет добавления навигационной системы на основе методов SLAM с использованием датчика ЛИДАР и камеры для ориентации в пространстве и взаимодействия с окружающей средой. Система будет развернута на базе одноплатного компьютера «Raspberry Pi4 Model B» с использованием ROS (Robot Operating System).

Литература:

1. Autonomous Warehouse Robots: a Brief History [Electronic resource] // inVia Robotics. – 2019, july. – URL: <https://www.inviarobotics.com/blog/autonomous-warehouse-robots-brief-history> (accessed 15.07.2024).
2. Moshayedi A.J., Li J., Liao L. AGV (automated guided vehicle) robot: Mission and obstacles in design and performance // Journal of Simulation & Analysis of Novel Technologies in Mechanical Engineering. – 2019. – Vol.12, N 4. – P. 0005–0018.
3. Каталог мотор редукторов для двигателя RS550 Far Along [Электронный ресурс]. – URL: https://aliexpress.ru/item/4000006871513.html?spm=a2g2w.orderdetail.0.0.629e4aa6IPuLzq&sku_id=12000026255850967 (дата обращения 05.05.2024).

«ИНТЕРНЕТ ВЕЩЕЙ (INTERNET OF THINGS — IOT): СТАНДАРТЫ, КОММУНИКАЦИОННЫЕ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, СЕТЕВЫЕ ПРИЛОЖЕНИЯ»

С. Галев, Д.Н. Климова

**Донской государственный технический университет,
г. Ростов-на-Дону, suhair.ghaleb1@gmail.com**

Научный руководитель: Климова Д.Н., к.п.н., доцент

В настоящее время Интернет вещей стал одной из важнейших современных технологий в мире высоких технологий. Благодаря быстрому развитию технологий и Интернета вещей Интернет вещей значительно развился, приобретая важное значение в различных потенциальных областях. Приложения IoT и связанные с ними исследовательские задачи. Например, «умные» города, здравоохранение, агрономия, логистика, розничная торговля - ожидается, что «умный» образ жизни и «умная» среда Интернета вещей проникнут практически во все аспекты повседневной жизни. Это делает его важной темой для исследований во многих странах. В данной исследовательской работе рассказывается об используемых стандартах, коммуникационных технологиях и сетевых приложениях.

Currently, the Internet of Things has become one of the most important modern technologies in the world of high technology. Due to the rapid development of technology and the Internet of Things, the Internet of Things has evolved significantly, becoming important in various potential areas. IoT applications and related research tasks. For example, smart cities, healthcare, agronomy, logistics, retail trade - it is expected that the "smart" lifestyle and the "smart" environment of the Internet of Things will penetrate into almost all aspects of everyday life. This makes it an important topic for research in many countries. This research paper describes the standards used, communication technologies and network applications.

Последние достижения в технологическом и научном развитии привели к появлению такого типа информационно-коммуникационных сетей, как Интернет вещей - это крупная сеть, занимающаяся передачей данных и информации, и Интернет вещей играет важную роль в развитии информационно-коммуникационных сетей, поскольку он представляет собой новую технологию для подключения отдельных устройств, а также устройств управления, будь то между ними в локальном масштабе в пределах одного здания или сооружения или в глобальном масштабе на уровне разных городов или стран по всему миру. Появлению Интернета вещей способствовали многие факторы, особенно в коммерческой и промышленной сферах, поскольку он предоставляет им достаточный объем информации, которую необходимо анализировать, обрабатывать и принимать практические

решения, помогающие им в различных сферах предпринимательства. Интернет вещей внес свой вклад в улучшение качества жизни населения и повлиял на их жизнь, облегчив процесс беспроводного обмена информацией путем подключения всех технических устройств через облачное соединение и передачи этих данных в электронное облако, их обработки и принятия решений без необходимости вмешательства человека. В результате Интернет вещей имеет набор критериев, в соответствии с которыми необходимо работать над созданием и развитием этих сетей и обменом информацией.

Анализ, проведенный за последние годы, показал нам количество достижений в области прикладных и информационных технологий и то, как это развитие оказало положительное влияние на жизнь общества.

1.Стандарты

Интернет вещей (IoT) включает в себя широкий спектр стандартов, которые охватывают основные термины безопасности Интернета вещей, архитектурные модели, интеграцию безопасности, классификацию средств защиты, протоколы безопасности, возможности подключения и взаимодействия между устройствами.

Стандарты, которые используются в ИОТ для подключенных к сети устройств приведены в таблице 1:

Стандарты	Описание	год
ГОСТ Р 70036-2022	Протокол беспроводной передачи данных на основе узкополосной модуляции радиосигнала (NB-Fi)	2022
ГОСТ Р ИСО/МЭК 29161-2019	Информационные технологии Структура Данных Уникальная идентификация для интернета вещей	2019
ПНСТ 877-2023	Информационные технологии. Интернет вещей. Методология обеспечения доверенности	2023
ПНСТ 921-2024	Информационные технологии. Интернет вещей	2024
ГОСТ Р 71168-	Интернет вещей Спецификация LoRaWAN RU	2024
ПНСТ 794-2022	Предварительный национальный стандарт российской федерации Информационные технологии Интернет вещей Периферийные вычисления	2022

Терминология, используемая в стандартах:

Стандарты для Интернета вещей (IoT) включают термины, относящиеся к протоколам безопасности, архитектурным моделям, методам интеграции и подключению устройств для обеспечения безопасного взаимодействия между устройствами.

уникальные идентификаторы устройств Интернета вещей, структуры данных и протоколы связи, облегчающие бесперебойное подключение и обмен данными. Кроме того, терминология в стандартах Интернета вещей охватывает такие аспекты, как протоколы беспроводной передачи данных, мониторинг окружающей среды и интеллектуальное городское планирование для повышения эффективности приложений Интернета вещей.

2. Интернет вещей в коммуникационные и информационные технологии, сетевые приложения

Интернет вещей (IoT) представляет собой большой трансформационный сдвиг в коммуникационных и информационных технологиях, помогая подключать устройства и обмениваться данными, а также способствуя развитию различных секторов, включая государственный и частный сектор. Интернет вещей и его различным приложениям во всех сферах жизни, таким как «Умные города», «Умная мобильность», «Умные дома», «Умное здравоохранение» и «умные сети». В «умных городах» Интернет вещей (IoT) основан на управлении дорожным движением посредством анализа данных в режиме реального времени, мониторинга окружающей среды и устойчивого городского планирования. В области здравоохранения Интернет вещей обеспечивает удалённое наблюдение за пациентами, использование носимых медицинских трекеров и систем управления приёмом лекарств, что произвело революцию в управлении здравоохранением. Кроме того, Интернет вещей расширяет возможности интеллектуальных сетей, обеспечивая точное управление энергопотреблением и интегрируя возобновляемые источники, Интернет вещей и его сети играют важную роль в демонстрации огромной роли в оказании значительного и эффективного воздействия на все сферы жизни.

Выводы:

Интернет вещей (IoT) произвёл революцию в различных секторах, таких как здравоохранение, «умные города» и промышленное управление, обеспечив удалённый мониторинг, точное управление энергопотреблением и интеграцию возобновляемых источников. Приложения Интернета вещей стали неотъемлемой частью

повседневной жизни, а интеллектуальная среда и образ жизни, как ожидается, проникнут во все сферы жизни общества, демонстрируя значительное влияние и быстрое развитие технологий Интернета вещей. Разработка стандартов для Интернета вещей, охватывающих протоколы безопасности, архитектурные модели и методы интеграции, подчеркивает важность безопасного подключения и взаимодействия между устройствами в меняющемся ландшафте Интернета вещей.

Литература:

1. *Kizza, J.M.* Internet of things (IoT): growth, challenges, and security. In: guide to computer network security. Texts in computer science // Springer, Cham, 2024.
2. *Rakhi Wajgi, Jitendra V. Tembhurne, Dipak W. Wajgi, Tapan Kumar Jain.* Internet of things // Springer Science+Business Media. 2023. – P. 131-157.
3. *Yachna, Gharde., A., K., Kamble., Swati, Dubey., Poornachandra.).* Smart Technologies for Smart Nations Internet of Things // International journal of innovative science and research technology, 2024. – P. 3762-3767.
4. *Yang-Hung Chang,* Application of Internet of Things technology in smart home system. – 2023. – Vol. 12645. – P. 126450J-126450J.

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ПОИСКА
ОБЪЕКТОВ МАЛОГО РАЗМЕРА НА ИЗОБРАЖЕНИЯХ**

Д.С. Апасова

**Новосибирский государственный технический университет,
г. Новосибирск, apasova.d@mail.ru**

Научный руководитель: Лихачев А.В., д.т.н., к.ф.-м.н.

В статье представлен сравнительный анализ методов поиска объектов малого размера на изображениях, что является ключевой задачей в области компьютерного зрения. Особое внимание уделяется важности предобработки изображений, которая включает такие этапы, как масштабирование, нормализация, устранение шума и выделение признаков. Рассматриваются популярные алгоритмы, включая метод опорных векторов (SVM), деревья решений и случайные леса, сверточные нейронные сети (CNN) и метод Хаара.

The article presents a comparative analysis of methods for searching for small objects in images, which is a key task in the field of computer vision. Special attention is paid to the importance of image preprocessing, which includes steps such as scaling, normalization, noise elimination and feature extraction. Popular algorithms are considered, including the support vector machine (SVM) method, decision trees and random forests, convolutional neural networks (CNN) and the Haar method.

Введение

Поиск объектов на изображениях представляет собой одну из ключевых задач в области компьютерного зрения. Эта задача приобретает особую актуальность в случаях, когда требуется обнаружить объекты малых размеров, такие как детали на производственной линии, мелкие элементы в медицинских изображениях или объекты в системах видеонаблюдения.

Эффективность алгоритмов поиска объектов зависит от множества факторов, включая качество исходных изображений, применяемые методы предобработки и архитектуру алгоритмов.

1. Важность предобработки изображений

Предобработка изображений – важный этап подготовки данных перед использованием в моделях машинного обучения. Предобработка изображений играет критическую роль в повышении точности и надежности алгоритмов поиска объектов. Она включает в себя ряд методов, направленных на улучшение качества изображений

перед их анализом. Основные этапы предобработки изображений включают:

Масштабирование — изменение размера изображений без потери качества.

Нормализация — приведение значений яркостей пикселей к определённому диапазону для улучшения сходимости и стабильности процесса обучения.

Цветовое преобразование — изменение фильтров изображения для адаптации к различным задачам.

Устранение шума и выделение контуров — удаление нежелательных артефактов и выделение границ объектов на изображении.

Выделение признаков — определение наиболее информативных особенностей изображения для дальнейшей классификации или распознавания.

Геометрические преобразования — изменение ориентации, поворота и положения изображения для адаптации к требованиям конкретной задачи.

Обрезка — удаление лишних частей изображения для улучшения качества или экономии ресурсов. [1]

Однако, набор и последовательность этапов зависит от решаемой задачи. Предобработка требуется в двух основных случаях:

Необходима очистка данных. Изображения могут быть сильно зашумлены, иметь искажённую перспективу и/или цвет, фон, не имеющий отношения к исследуемому объекту;

Необходима аугментация данных. Аугментация изображений – это методика создания дополнительных данных из имеющегося набора. В условиях небольшого набора данных аугментация позволит улучшить результаты обучения модели.

2. Популярные методы поиска объектов

Существует множество алгоритмов и подходов для поиска объектов на изображениях. Рассмотрим несколько наиболее популярных методов, используемых для обнаружения мелких объектов.

2.1 Метод опорных векторов (SVM)

Метод опорных векторов — это алгоритм машинного обучения, применяемый для задач линейной и нелинейной классификации, регрессии и обнаружения аномальных данных.[2] Чтобы понять, как работает алгоритм, представим, что у нас есть набор точек на бумаге и эти точки принадлежат двум разным группам. Например, это картинки с кошками и собаками. SVM помогает нарисовать линию (или, в более сложных случаях, плоскость или гиперплоскость), которая максимально

четко разделяет эти две группы точек. В случае с животными алгоритм, используя характеристики — размер ушей, форма хвоста и т. д., — разделит изображения так, что все кошки окажутся с одной стороны, а все собаки — с другой.

2.2 Деревья решений и случайные леса

Деревья решений и случайные леса являются популярными методами для классификации изображений. Дерево решений — это древовидная модель, которая подсказывает, в каком порядке проверять характеристики объекта, чтобы получить его метку.[3] Внутренние узлы сообщают, какие признаки нужно проверить, а листья показывают, какой прогноз даёт дерево. Прогнозы делаются следующим образом. Каждый лист содержит подмножество обучающего набора данных. Все его экземпляры проходят все проверки на пути от корня к листу. Экземпляры — это отдельные объекты или наблюдения в обучающем наборе данных. Они проходят через дерево, начиная с корня, и проверяются на соответствие условиям, заданным внутренними узлами. Если экземпляр проходит все проверки, он попадает в лист, где делается прогноз. При прогнозировании результата для нового объекта мы присваиваем ему объединённую метку обучающих данных, которые заканчиваются на том же листе, что и экземпляр. Деревья решений страдают от нескольких проблем. Они склонны к переобучению. Поскольку точность повышается с каждым внутренним узлом, при обучении, дерево будет стремиться к максимальному размеру, чтобы улучшить показатели производительности. Это ухудшает способность дерева к обобщению и его эффективность при работе с невидимыми данными. Невидимые данные — это те данные, которые дерево не видело во время обучения. Ухудшение способности дерева к обобщению и его эффективности при работе с невидимыми данными происходит потому, что дерево начинает моделировать шум вместо важных закономерностей, что приводит к снижению качества предсказаний, поскольку оно начнёт моделировать шум. Можно заранее ограничить глубину дерева, но проблема неустойчивости всё равно остаётся. А именно, даже небольшие изменения в обучающих данных, такие как исключение нескольких примеров, могут привести к совершенно другому дереву. Таким образом, возникает вопрос, какому дереву доверять. Случайные леса решают обе проблемы. Идея состоит в том, чтобы построить несколько деревьев, используя разные подмножества обучающих данных и признаков для каждого дерева в лесу. Затем их прогнозы объединяются, например, выводится результат голосования большинства или среднее значение.

2.3 Свёрточные нейронные сети (CNN)

Свёрточные нейронные сети — разновидность нейросетей для обработки данных с сеточной структурой: изображений и видео. Такие нейросети анализируют пиксели, которые находятся близко друг к другу и содержат непрерывную визуальную информацию — яркость и оттенок. Например, если в одном пикселе нейросеть видит цветок, то и в стоящих рядом пикселях она тоже его распознаёт. По своей структуре сверточные нейронные сети – это воронка: все начинается с картины в целом, затем внимание переключается на отдельные детали. Мозг работает точно так же: человек на улице видит сначала кошку и лишь потом начинает рассматривать цвет ее шерсти и глаз.

2.4 Метод Хаара

Метод Хаара — это классический подход к обнаружению объектов, основанный на обучении каскадных классификаторов. Каскадная функция обучается на основе множества положительных и отрицательных изображений. Характеристики Хаара представляют собой простые прямоугольные шаблоны, которые могут быть использованы для извлечения признаков из изображения. Они могут быть представлены в виде белых и черных прямоугольников, которые указывают на области, где вычисляется разность яркости. Например, одна из характеристик может определять разницу между средним значением яркости верхней и нижней частей изображения. Метод Хаара использует каскад классификаторов для повышения производительности. Каждый классификатор в каскаде принимает решение о том, содержит ли текущее окно изображения искомый объект. Если классификатор принимает негативное решение, окно отбрасывается, и процесс продолжается только для тех окон, которые прошли проверку. Это позволяет значительно сократить количество окон, которые нужно обрабатывать, и ускоряет процесс детектирования.[4]

Заключение

Поиск объектов малого размера на изображениях представляет собой сложную задачу, требующую тщательного подхода к выбору методов и предварительной обработки изображений. Эффективные методы предварительной обработки играют ключевую роль в повышении точности и надёжности алгоритмов. Современные методы глубокого обучения, такие как свёрточные нейронные сети, демонстрируют впечатляющие результаты, однако классические методы остаются актуальными в определённых приложениях благодаря своей простоте и скорости. Важно учитывать специфические требования задачи и условия, в которых будет применяться алгоритм, чтобы выбрать наиболее подходящий метод поиска объектов.

Литература:

1. *Русакова Е.И.* Основные шаги по подготовке изображений для обучения систем компьютерного зрения / *Русакова Е. И.* [Электронный ресурс] // ВШЭ. – URL: https://perm.hse.ru/data/2023/06/08/2020097691/Семинар_12_05_НУТ.pdf (дата обращения: 29.10.2024).
2. *Останина А.* Support Vector Machine: классификация данных с помощью метода опорных векторов / *Останина А.* [Электронный ресурс] // Skillfactory. – URL: <https://blog.skillfactory.ru/svm-metod-opornyh-vektorov/> (дата обращения: 30.10.2024).
3. *Milos Simic, Michal Aibin.* Decision Trees vs. Random Forests / *Milos Simic, Michal Aibin* [Электронный ресурс] // Baeldung. – URL: <https://www.baeldung.com/cs/decision-trees-vs-random-forests> (дата обращения: 31.11.2024).
4. *Амеличев Г. Э., Панина В. С., Белов Ю. С.* Распознавание лиц с использованием каскадов Хаара / *Амеличев Г. Э., Панина В. С., Белов Ю. С.* [Электронный ресурс] // Cyberleninka. – URL: <https://cyberleninka.ru/article/n/raspoznavanie-lits-s-ispolzovaniem-kaskadov-haara> (дата обращения: 01.11.2024).

РИСКИ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРОЕКТАХ, СВЯЗАННЫХ С ГЕНЕРАЦИЕЙ ВИЗУАЛЬНОГО КОНТЕНТА

М.С. Давыдов

Новосибирский государственный технический университет,

г. Новосибирск, TSMFDR@yandex.ru

Научный руководитель: Вакорин М.П., к.э.н., доцент

В этом тезисе описываются возможности и потенциальные риски применения искусственного интеллекта в различных сферах, связанных с созданием визуального контента, в частности в маркетинге, архитектуре, кинематографе и новостях. После выявления рисков в выбранных сферах сделана классификация рисков применения искусственного интеллекта в данных сферах.

This thesis describes the opportunities and potential risks of applying artificial intelligence in various fields related to the creation of visual content, in particular in marketing, architecture, cinematography and news. After identifying the risks in the selected areas, a classification of the risks of applying artificial intelligence in these areas is made.

С приходом современных высокоразвитых моделей искусственного интеллекта, его использование в различных проектах, будь то социальные, научные или организационные, экспоненциально увеличилось. Практически невозможно найти проект, где не используется искусственный интеллект, даже в малой степени. Не смотря на удобство, которое принес этот инструмент, у искусственного интеллекта есть масса рисков, которые требуют изучения.

Риск — это потенциальное событие, которое может оказать как положительное, так и отрицательное воздействие на успех проекта. Риски могут быть вызваны внешними факторами (например, изменениями рынка) или внутренними причинами (например, ошибками в планировании). Они могут привести как к дополнительным возможностям, так и к проблемам, и возникают на всех этапах жизненного цикла проекта [1].

Искусственный интеллект – это широкое понятие, охватывающее различные технологии, позволяющие интеллектуальным агентам выполнять задачи, требующие интеллектуальных усилий, имитируя некоторые возможности человеческого мозга. ИИ может обучаться на основе данных, распознавать образы, понимать естественный язык и многое другое. Существует множество классификаций ИИ, например, по используемым методам (символьный подход, нейронные сети, эволюционные алгоритмы) или по функциональности (реактивные машины, системы с ограниченной памятью) [2].

Искусственный интеллект в наши дни используется повсеместно: в медицине, играх, математике, финансах, генерации изображений и др. Для генерации контента чаще всего используются нейронные сети, способные к обучению на примерах.

Прорыв в сфере создания визуального контента, вызванный развитием искусственного интеллекта, открыл перед дизайнерами, маркетологами и другими специалистами новые возможности и перспективы. Однако появились и определенные риски, которые необходимо учитывать.

В маркетинге генерация изображений с помощью искусственного интеллекта повсеместна. Например, на данный момент 61% B2B-маркетологов в США используют ИИ для создания чат-ботов, кодирования, дизайна и контента. С помощью ИИ маркетологи создают рекламные баннеры, дизайны для брендов, оптимизируют рекламные компании, автоматизируют процессы, улучшение прогнозирования поведения клиентов [3].

Внедрение ИИ в маркетинг требует значительных инвестиций, как финансовых, так и временных. Кроме того, необходимы специалисты,

обладающие знаниями в области ИИ и машинного обучения, в рыночных условиях их дефицита. Одной из важных проблем внедрения ИИ также является отсутствие законодательского регулирования и стандартов, из-за чего появляется сложность реализации и оценки финального результата. Так же одним из рисков внедрения является проблематика конфиденциальности. Сбор и анализ больших объемов данных о клиентах поднимают вопросы о безопасности данных и их нецелесообразным использованием [4][5].

В архитектуре использование генеративные модели искусственного интеллекта позволяют быстрее создавать реалистичные 3D-модели зданий и визуализировать проекты в различных сценариях [6][7].

Однако внедрить ИИ можно не во все области архитектурного проектирования. Также внедрение ИИ совместимо не со всеми существующими инструментами и процессами. Архитекторам и строителям приходится осваивать новые навыки и знания в области искусственного интеллекта, чтобы эффективно использовать его инструменты и понимать их ограничения. Сбор, очистка и анализ больших объемов данных, необходимых для эффективной работы ИИ в архитектурных проектах, требуют значительных временных и финансовых ресурсов. Малая разработанность правовых норм и стандартов, регулирующих применение ИИ в архитектуре и градостроительстве, создает дополнительные сложности и риски, в частности этического характера. К тому же, некоторые архитекторы высказали, что с помощью ИИ невозможно создать абсолютно новые дизайнерские решения [6][7].

В области кинематографа ИИ используется для генерации визуальных эффектов и написания сценариев. Также используется для анализа трейлеров, реакции аудитории и зрительских предпочтений. Существуют и примеры фильмов, полностью созданным искусственным интеллектом [8].

Использование искусственного интеллекта в кинематографе несет в себе и существенные риски. Одной из главных опасений является возможность массовой безработицы среди специалистов, таких как актеры, сценаристы и художники по визуальным эффектам, что приводит к массовым забастовкам. Кроме того, недостаточное понимание принципов работы ИИ среди широкого круга специалистов в условиях постоянного производства фильмов, где специалисты нередко работают над несколькими проектами сразу, приводит к снижению качества фильмов и критики со стороны зрителей. Также имеются этические проблемы использования внешности скончавшихся актеров и создания их цифровых клонов. В индустрии развлечений актеры

озвучивания, восходящие таланты, умершие знаменитости и люди, создавшие контент в прошлом, могут стать жертвами недобросовестного использования ИИ без их согласия и существующие законы не могут их от этого защитить [9].

В сфере новостей ИИ используется для автоматизации сбора информации и некоторых процессов публикации новостей. Однако имеются существенные опасения распространения фальшивых новостей, дезинформации, так как полученная информация недостаточно фильтруется. Также есть риски, связанные с распространением социальных предубеждений в новостных СМИ. Существует несчетное количество случаев, когда дипфейки, фотореалистичное объединения изображений с помощью ИИ, были представлены как факты [10][11]. Нейросети также способны генерировать визуальный контент, противоречащий общепринятым нормам межрелигиозных и межрасовых отношений [12].

Использование развитых генеративных моделей искусственного интеллекта влечет за собой масштабные сокращения рабочих мест. Так, в мае 2023 года около 4000 человек потеряли работу из-за внедрения ИИ. Согласно некоторым прогнозам, к 2035 году число потерявших работу из-за ИИ может затронуть до 300 миллионов человек [13].

Подытожив, можно сформировать классификацию рисков использования искусственного интеллекта в проекты генерации визуального контента, которая представлена в таблице 1.

Таблица 1 – Классификация рисков

Тип риска	Подробности
Этические риски	Использование данных, полученных без согласия, сокращение рабочих мест, генерация контента, который может относиться к клевете, утрата аутентичности, генерация контента, нарушающего религиозные, этические нормы и т.д.
Экономические риски	Переобучение специалистов, выделение больших бюджетов, риск монополии, потеря рабочих мест и т.д.
Технические риски	Низкое качество исходных данных и необходимость их перебора и анализа, сложность определения, создан ли контент человеком или ИИ, и возможная киберуязвимость.

Правовые риски	Отсутствие правовых норм и стандартов, регулирующих применение ИИ, возможное нарушение авторских прав и конфиденциальности персональных данных при получении исходных данных.
----------------	---

По результатам проведенного исследования были выявлены потенциальные риски применения искусственного интеллекта в различных сферах, связанных с созданием визуального контента. ИИ имеет множество преимуществ, но также много рисков, которые стоит учитывать при внедрении его в проект.

Литература:

1. ГОСТ Р 51897-2021. Национальный стандарт российской федерации. Менеджмент риска. Термины и определения. – Введ. 2022–03–01. – М.: Российский институт стандартизации, 2021. – 20 с.
2. *Рассел Стюарт*. Искусственный интеллект: современный подход / Стюарт Рассел, Питер Норвиг. — 4-е. — Хобокен: Пирсон, 2021.
3. *Savage M.* AI And Marketing: To Lean Into Or Lean Away? [Электронный ресурс] // Forbes Innovation – 2024, March – URL: <https://www.forbes.com/sites/sap/2024/03/21/ai-and-marketing-to-lean-into-or-lean-away/?sh=2678f8f06713> (accessed 07.10.2024).
4. *Гребенькова Д., Новикова Е.* Искусственный интеллект: возможности и риски для маркетинга и бизнеса. [Электронный ресурс] // New Retail – № 3. – 2024 – URL: https://new-retail.ru/tehnologii/iskusstvennyy_intellekt_vozmozhnosti_i_riski_dlya_marketinga_i_biznesa/ (дата обращения: 05.10.2024).
5. *Рябая В.С., Голик В.С.* Искусственный интеллект в маркетинге как составляющая развития инновационной экономики //Science and innovation. – 2024. – Т. 3. – №. Special Issue 24. – С. 446-450.
6. *Гайлитис В.С., Азаренко Н.Ю.* Искусственный интеллект в архитектуре //ББК 32.81: 32.973 Ц 75 Редакционная коллегия: Глав. ред. к. э. н., доц. Казаков ОД, отв. ред. к. э. н., доц. Азаренко НЮ, к.э.н., доц. Юркова ОН. – 2024. – С. 163.
7. *Власова Е.Л. и др.* Искусственный интеллект в архитектурно-градостроительном проектировании //Architecture and Modern Information Technologies. – 2023. – №. 4 (65). – С. 311-324.
8. *Сибилев В.О.* Применение искусственного интеллекта в сфере кинематографа при создании фильма //ББК – 2024 – №34. – С. 3494-3500.

9. Carthy C.R. Artificial Intelligence and the Entertainment Industry. // Honors Theses and Capstones. – 2024 – С. 857.
10. Dhiman D.B. Does Artificial Intelligence Help Journalists: A Boon or Bane? //SSRN 4401194. – 2023.
11. Yijing X. The Application of Artificial Intelligence in News Communication: Practical Exploration and Ethical Dilemma //Media Studies. – 2024. – Т. 1. – №. 2. – С. 28-35.
12. Barroso da Silveira J., Alves Lima E. Racial Biases in AIs and Gemini's Inability to Write Narratives About Black People //Emerging Media. – 2024. – С. 27523543241277564.
13. Вечерова Е. Дизайнеры, чиновники, маркетологи: кто еще может скоро потерять работу из-за ИИ [Электронный ресурс] // Forbes. Технологии: – № 6. – 2023. – URL: <https://www.forbes.ru/tekhnologii/491761-dizajnery-cinovniki-marketologi-k-to-ese-mozet-skoro-poterat-rabotu-iz-za-ii?ysclid=m1z86bxo1x13648972>

АВТОМАТИЗАЦИЯ УЧЕТНОЙ ДЕЯТЕЛЬНОСТИ ЧАСТНОЙ МЕДИЦИНСКОЙ КОМПАНИИ

Д.М. Захаров

**Донской государственный технический университет
г. Ростов-на-Дону, zach.dan.m@gmail.com**

Научный руководитель: Климова Д.Н., к.т.н., доцент

Автоматизация учетной деятельности в частной медицинской компании – это процесс внедрения информационных технологий для оптимизации и ускорения рутинных операций. Статья рассматривает преимущества автоматизации, такие как повышение точности данных, сокращение времени на обработку информации и улучшение качества обслуживания пациентов. Результатом работы является готовое приложение для автоматизации учета частной медицинской компании. Целевая аудитория: руководители медицинских учреждений, финансовые директора, IT-специалисты.

Automation of accounting activities in a private healthcare company is the process of implementing information technology to optimise and speed up routine operations. This article discusses the benefits of automation, such as increased data accuracy, reduced processing time and improved patient care. The result of work is the ready application for automation of accounting of private medical company. Target audience: heads of medical institutions, financial directors, IT-specialists.

Введение

В современном мире автоматизация процессов играет ключевую роль в повышении эффективности и качества услуг, предоставляемых

медицинскими учреждениями. Особенно это актуально для платных медицинских центров, которые сталкиваются с рядом специфических задач и требований. В данном исследовании рассмотрим важность автоматизации процессов в контексте платных медицинских учреждений. Будут рассмотрены основные аспекты, включая оптимизацию рабочего времени персонала, улучшение качества обслуживания пациентов, контроль финансовых операций и обеспечение безопасности данных. Такой подход не только повышает эффективность работы учреждения, но и способствует созданию более благоприятной среды для оказания медицинской помощи и повышению уровня удовлетворенности клиентов.

В настоящее время непрерывно развивается платформа 1С, занимая растущий сегмент рынка информационных услуг в сфере автоматизации различных бизнес-процессов. Использование платформы 1С для автоматизации работы в медицинских учреждениях имеет свои плюсы. Во-первых, она предоставляет обширный набор инструментов, с помощью которых можно управлять практически всеми аспектами работы клиники, начиная от учета пациентов и заканчивая финансовым учетом. Кроме того, 1С обладает гибкостью: систему можно настроить под индивидуальные потребности и процессы каждой конкретной клиники, что делает ее работу более удобной и эффективной. Также важно отметить поддержку и помощь, доступные при использовании 1С: система широко используется, и поэтому помощь и поддержка всегда доступны, если возникнут какие-либо проблемы. Интеграция 1С с другими системами также обеспечивает беспрепятственное взаимодействие различных частей медицинского учреждения. И наконец, использование 1С может оказаться более экономически эффективным, чем разработка собственной системы с нуля, благодаря широкому распространению и поддержке платформы.

Данная работа направлена на создание конфигурации для платного медицинского учреждения, с учетом определенных запросов и предпочтений для данного учреждения. Такое решение позволит максимально адаптировать систему автоматизации под конкретные потребности и особенности работы медицинского учреждения. Создание специализированной конфигурации обеспечит оптимизацию процессов внутри учреждения и повышение эффективности работы персонала. Благодаря индивидуальному подходу к разработке, система будет более гибкой и лучше адаптированной к изменениям внутри и вне учреждения. Это поспособствует повышению удовлетворенности клиентов и качества обслуживания в медицинском учреждении, делая его работу более эффективной и успешной.

Объект исследования: частное медицинское учреждение.

Предмет исследования: комплекс мероприятий по разработке конфигурации 1С для платного медицинского учреждения.

Цель работы: реализация конфигурации 1С для автоматизации работы платного медицинского учреждения.

Задачи для выполнения поставленной цели:

- Анализ предметной области.
- Изучение существующих решений.
- Анализ требований к разрабатываемой системе автоматизации учета.
- Программная реализация приложения.
- Пользовательский интерфейс.
- Экономическое обоснование процесса внедрения программного продукта.

Предметно-ориентированная среда разработки, внедренная в платформе «1С: Предприятие 8.3», доступна в составе конфигуратора, который представляет собой инструментарий для создания и модификации конфигураций. Разработка приложений ведется с использованием языка программирования 1С.

Теоретическая значимость работы: решение проблемы автоматизации учетной деятельности платного медицинского учреждения.

Практическая значимость работы: реализация пользовательского интерфейса для платного медицинского учреждения на платформе «1С: Предприятие 8.3».

Анализ предметной области

В медицинской индустрии платные медицинские учреждения занимают важное место, предоставляя широкий спектр медицинских услуг за плату. Они отличаются своей коммерческой направленностью, в отличие от государственных учреждений, и в то же время, у них есть определенные особенности, которые делают их уникальными в контексте организации учетной деятельности [1].

Коммерческая направленность. В отличие от государственных медицинских учреждений, платные клиники и больницы ориентированы на коммерческую деятельность. Это означает, что их основная цель - получение прибыли от предоставления медицинских услуг. Как бизнесы, платные медицинские учреждения должны учитывать финансовые аспекты своей деятельности, такие как расходы на оборудование и персонал, ценообразование на услуги, а также управление доходами и расходами [2].

Автоматизация учетной деятельности в платных медицинских учреждениях является ключевым аспектом их успешной работы и эффективного управления. В современном медицинском мире, где информация играет решающую роль, эффективное управление данными и финансами становится необходимостью [3].

Для выполнения всех поставленных задач, мною было принято решение выбрать среду 1С [4, 5] по ряду следующих причин:

Многофункциональность:

- 1С прекрасно адаптируется под конкретные требования бизнеса, будь то учет и финансы, управление ресурсами предприятия, управление персоналом, торговля, продажи.

- Управление взаимоотношениями с клиентами (CRM): некоторые версии 1С включают в себя функционал CRM, который позволяет управлять взаимоотношениями с клиентами, ведение базы данных клиентов, анализ продаж и маркетинговых активностей.

- Аналитика и отчетность: 1С обеспечивает возможность анализа данных и формирования отчетов для принятия управленческих решений. Пользователи могут создавать различные виды отчетов, дашбордов и аналитических обзоров для мониторинга бизнес-процессов и оценки их эффективности.

- Интеграция с другими системами: 1С обладает возможностью интеграции с другими системами, такими как банковские системы, системы электронной коммерции, системы управления взаимоотношениями с клиентами и т. д., что позволяет создать единое информационное пространство и обеспечить взаимодействие между различными подсистемами предприятия.

Высокая эффективность внедрения:

- Простота использования и настройки: Интерфейс 1С интуитивно понятен и легко настраивается под конкретные потребности организации. Это позволяет быстро освоить систему и адаптировать ее к специфике бизнес-процессов.

- Готовые решения и шаблоны: 1С предоставляет готовые шаблоны и конфигурации для различных отраслей и видов деятельности. Это упрощает процесс внедрения, поскольку значительная часть функциональности уже реализована и готова к использованию.

- Обширное сообщество разработчиков и консультантов: 1С имеет широкое сообщество разработчиков, консультантов и пользователей, готовых поделиться опытом, советами и решениями. Это обеспечивает доступ к профессиональной поддержке и помощи в решении возникающих проблем.

- Быстрая разработка и внедрение: благодаря готовым шаблонам и конфигурациям, а также простоте настройки и использования, разработка и внедрение систем на базе 1С происходит значительно быстрее по сравнению с созданием систем с нуля.

- Надежность и стабильность: 1С известна своей надежностью и стабильностью работы. Это позволяет уверенно использовать систему в повседневной деятельности организации, минимизируя риски возникновения сбоев и проблем.

- Поддержка и обновления: 1С регулярно выпускает обновления и новые версии своего программного обеспечения, обеспечивая поддержку и современные возможности для пользователей. Это позволяет поддерживать систему в актуальном состоянии и использовать последние технологические достижения.

Простота и удобство работы с 1С:

- Интуитивный интерфейс: 1С обладает интуитивно понятным интерфейсом, который позволяет пользователям быстро освоить основные функции системы без необходимости специального обучения или опыта в работе с аналогичными программами.

- Гибкость и настраиваемость: Пользователи могут легко настраивать интерфейс и функционал системы под свои индивидуальные потребности и предпочтения. Это позволяет создавать персонализированные рабочие пространства и оптимизировать процессы работы.

- Готовые шаблоны и конфигурации: 1С предоставляет готовые шаблоны и конфигурации для различных видов деятельности и отраслей, что упрощает процесс начальной настройки и внедрения системы. Пользователи могут выбрать подходящий шаблон и быстро адаптировать его под свои нужды.

- Простота в использовании функционала: Функционал 1С представлен в виде интуитивно понятных и легко доступных инструментов и возможностей, что позволяет пользователям быстро выполнять необходимые операции и задачи без лишних усилий и сложностей.

- Эффективность работы: благодаря простоте и удобству использования, 1С позволяет пользователям эффективно выполнять свои задачи и повышать производительность труда без необходимости в длительном обучении и адаптации к системе.

Изучение имеющихся решений

«1С: Медицина» — это программное решение, разработанное на базе платформы «1С: Предприятие», специально для автоматизации учета и

управления бизнес-процессами в медицинских учреждениях. Это комплексное программное обеспечение предоставляет широкий спектр функциональных возможностей, которые включают в себя учет пациентов, ведение медицинской документации, запись на прием, управление расписанием врачей, финансовый учет, аналитику и отчетность [6, 7]. «1С: Медицина» представляет собой мощное программное решение для автоматизации медицинских учреждений, но оно имеет ряд аспектов, которые никак не подойдут для нашего платного медицинского учреждения.

«DocaCRM» — это облачная система управления клиникой, разработанное для автоматизации работы медицинских учреждений, включая частные клиники и другие организации здравоохранения. В своей основе «DocaCRM» предоставляет набор инструментов для управления взаимодействием с клиентами (CRM) с учетом специфики медицинской отрасли [8]. В целом, хотя «DocaCRM» имеет ряд преимуществ для управления медицинским учреждением, существуют ряд факторов, которые делают его неподходящим для нашего случая. Далее приведены рисунки, показывающие программную составляющую приложений

Анализ требований к разрабатываемой системе автоматизации учета

При разработке конфигурации 1С для платного медицинского учреждения были учтаны функциональные требования, учреждения для максимально эффективного управления и автоматизации его ключевых бизнес-процессов [9, 10]. Ниже представлены основные компоненты и функции, которые должны будут реализованы в конфигурации:

Учет пациентов и клиентов:

- регистрация и ведение базы данных пациентов с полной персональной информацией и медицинской историей,
- возможность быстрого поиска и фильтрации информации о пациентах для удобства работы персонала.

Учет медицинских услуг:

- учет и категоризация медицинских услуг с указанием тарифов и стоимости,
- формирование счетов на оплату медицинских услуг для пациентов и страховых компаний.

Учет финансов и финансовые операции:

- ведение учета финансовых операций, включая выставление счетов, оплату услуг, учет доходов и расходов учреждения,

– формирование отчетов о финансовой деятельности и анализе финансовых показателей.

Возможность анализа эффективности работы учреждения и принятия управленческих решений на основе полученных данных.

Выводы

Исследование подчеркивает важность разработки данного программного продукта. Оно позволит осуществлять процессы ведения учета частной медицинской компании: учет пациентов и клиентов, учет медицинских услуг, учет финансов и финансовых операций, управление складом и медицинским оборудованием. Все это позволит упростить ведение учета, сократить время затраты на рутинные задачи компании, быстрее получать важные сведения для принятия управленческих решений.

Литература

1. *Радченко, М.Г.* 1С: Предприятие 8.3. Практическое пособие разработчика. Примеры и типовые приемы / М.Г. Радченко, Е.Ю. Хрусталева. – М.: 1С-Паблишинг, 2011. - 965 с.

2. *Хрусталева Е.Ю., Радченко М.Г.* Язык запросов «1С: Предприятия 8» / Е.Ю. Хрусталева, Радченко М. Г.: 000 «1 С-Паблишинг», 2013. – 358 с.

3. 1С: Предприятие 8.3: комьюнити-лицензия [для разработки и отладки прикладных решений (конфигураций) на базе технологической платформы 1С:Предприятие] / разработчик «1С». – Москва. – Электронная программа: электронная.

4. *Бартенев, О.В.* 1С: Предприятие: программирование для всех / О.В. Бартенев. - М.: Диалог-Мифи, 2016. - 464 с.

5. *Кашаев С.М.* Программирование в 1С: Предприятие 8.3. СПб.; Питер, 2014. – 304с.

6. Хакасия.ру – официальный сайт. – Москва. – Обновляется автоматически. – URL: 1С: Медицина (khakasia.ru) (дата обращения: 29.10.2024).

7. Miterra.ru – официальный сайт. – Москва. – Обновляется в течении дня. – URL: 1С: Медицина. Поликлиника | miterra.ru (дата обращения: 29.10.2024).

8. Doca.ru – официальный сайт. – Набережные Челны. – Обновляется в течении дня. – URL: Главная - DOCA - облачная система управления клиникой (doca.ru) (дата обращения: 01.11.2024).

9. *Гончаров Д.И., Хрусталев Е.Ю.* Технологии интеграции «1С: Предприятие 8.2». – 1С-Паблишинг, 2011. – 358 с.

10. Митичкин С.А. Разработка в системе 1С: Предприятие 8.3. – Спб.: «1С». 2017-384с.

АНАЛИЗ МЕТОДОВ РАСПОЗНАВАНИЯ ЖЕСТОВОГО ЯЗЫКА С ИСПОЛЬЗОВАНИЕМ НЕМАНУАЛЬНЫХ КОМПОНЕНТОВ

Д.С. Константинов

**Новосибирский государственный технический университет,
г. Новосибирск, d7hdmk@gmail.com**

Научный руководитель: Гриф М.Г., к. т. н., профессор АСУ

В данной работе рассматривается актуальность распознавания русского жестового языка, понятие немануального компонента, а также сравниваются современные модели распознавания жестовых языков, которые учитывают немануальный компонент.

This paper examines the relevance of Russian sign language recognition, the concept of a non-manual component, and compares modern models of sign language recognition that take into account the non-manual component.

Автоматизация распознавания жестовых языков играет важную роль в контексте взаимодействия слышащих людей с людьми, использующими жесты в качестве способа коммуникации. По данным Всемирной федерации глухих в России более 13 миллионов людей с нарушением слуха [1]. Из этой статистики можно сделать вывод, что в России проблема взаимодействия слышащих людей с глухими людьми, использующими РЖЯ, является актуальной.

РЖЯ состоит из жестов, означающих слова, буквы, цифры и т.д. Жест состоит из нескольких компонент, таких как:

- 1) конфигурация;
- 2) место исполнения жеста;
- 3) направление движения;
- 4) характер движения;
- 5) немануальный компонент [2].

Немануальный компонент – это элементы жеста, которые не связаны с руками, например, положение головы, выражение лица, мимика.

В РЖЯ в каждом жесте присутствует немануальный компонент, но существуют так называемые комбинированные жесты, для которых немануальный компонент является обязательным, а также омонимы – жесты, которые можно отличить друг от друга только по немануальному компоненту [3].

Данный компонент является одним из важных составляющих РЖЯ.

В современных исследованиях распознавания РЖЯ немануальному компоненту уделяется малое количество времени, либо он игнорируется. Но использование немануального компонента при распознавании может увеличить качество распознавания РЖЯ.

Всё это даёт понимание, что задача распознавания немануальных компонент РЖЯ является актуальной.

Цель данной статьи — анализ методов, применяемых для распознавания немануальных компонентов в рамках задачи автоматического распознавания жестовых языков.

Модели распознавания РЖЯ делятся на два вида:

- 1) изолированные – распознаётся отдельно взятый жест;
- 2) непрерывные – распознаётся не отдельный жест, а поток жестов.

С практической точки зрения более важна задача по непрерывному распознаванию жестов, так как речь человека происходит непрерывно.

На данный момент не существует моделей для распознавания РЖЯ, которые в полной мере используют немануальные компоненты.

В зарубежных источниках были найдены следующие современные модели непрерывного распознавания иностранных жестовых языков, которые учитывают немануальные компоненты:

- 1) BiLTSM;
- 2) MNM-SLR;
- 3) MM-SLR.

Первая модель использует три архитектуры BiLTSM объединённых в одну [4].

Вторая модель использует в своей архитектуре двухпоточную свёрточную сеть, первый поток использует VGG16 для извлечения немануального компонента, второй распознаёт жест по руке и далее потоки соединяются при помощи BiLTSM [5].

Третья модель использует в своей архитектуре YOLOv5 для извлечения областей лица и рук, C3D для извлечения признаков в различных модальностях, а также сеть LSTM для получения пространственных и временных дескрипторов для классификации [6].

Были сопоставлены результаты точности распознавания данных моделей. Тестирование производилось на наборе данных RWTH-PHOENIX. Точность вычислялась по формуле (TP – количество правильно обнаруженных жестов, FP – количество ошибочно обнаруженных жестов):

$$Accuracy = \frac{TP}{TP + FP}, \quad (1)$$

Результаты представлены в таблице 1.

Таблица 1 – Результаты точности моделей

Название	Точность, %
BiLTSM	73.99
MNM-SLR	94.37
MM-SLR	82.5

Из таблицы 1 можно сделать вывод, что модель MNM-SLR с распознаванием иностранного жестового языка справляется лучше. Однако необходимо отметить, что в рамках распознавания РЖЯ модели могут показать другие результаты.

В рамках дальнейшей работы планируется сформировать набор данных для немануальных компонентов РЖЯ, произвести дообучение представленных моделей и выявить слабые стороны для их устранения.

Литература:

1. *Варинова О.А.* Особенности описания социального сообщества глухих людей // Теория и практика общественного развития. – 2020. – №. 7 (149). – С. 44-48.
2. *Давиденко Т., Комарова А.А.* Краткий очерк по лингвистике русского жестового языка // Москва. – 2006.
3. Введение в лингвистику жестовых языков. Русский жестовый язык: Учебник / С.И. Буркова, В. И. Киммельман, Е. В. Филимонова [и др.]. – Новосибирск: Новосибирский государственный технический университет, 2021. – 356 с. – (Учебники НГТУ).
4. *Sánchez-Ruiz D., Olvera-López J. A., Olmos-Pineda I.* Sign Language Recognition through Manual and Non-Manual Features // Research in Computing Science. – 2022. – Т. 150. – №. 12. – С. 131–141.
5. *Jebali M., Dakhli A., Bakari W.* Deep learning-based sign language recognition system using both manual and non-manual components fusion // AIMS Math. – 2024. – Т. 9. – С. 2105-2122.
6. *Javaid S., Rizvi S.* Manual and non-manual sign language recognition framework using hybrid deep learning techniques // Journal of Intelligent & Fuzzy Systems. – 2023. – Т. 45. – №. 3. – С. 3823-3833.

АВТОМАТИЗАЦИЯ АУДИТА БЕЗОПАСНОСТИ ОС АЛЪТ

В.Е. Лесик, Д.В. Никифоров
РГУ нефти и газа (НИУ) имени И.М. Губкина,
г. Москва, nikiforov2020@inbox.ru
Научный руководитель: Павловский В.В.

Информационно-вычислительная система представляет собой комплекс баз данных, систем управления ими и прикладных программ, работающих как единое целое на вычислительных ресурсах для решения конкретных задач, согласно [1]. В рамках данной статьи понятие «информационная система» охватывает информационно-вычислительные системы, операционные системы, базы данных и сетевые устройства [2]. Современные стандарты, такие как стандарты PCI DSS и рекомендации ФСТЭК, предъявляют строгие требования к безопасности конфигурации и управлению уязвимостями [3-4].

An information-computing system is a complex of databases, their management systems and application programmes working as a whole on computing resources to solve specific tasks, according to [1]. For the purpose of this paper, the term 'information system' encompasses information and computing systems, operating systems, databases, and network devices [2]. Modern standards, such as PCI DSS standards and FSTEC recommendations, impose strict requirements for configuration security and vulnerability management [3-4].

Для оценки соответствия этим требованиям необходим аудит безопасности. Однако сложность и динамика информационных систем делают выполнение периодических и полных проверок крайне трудоемким. Особенно это заметно на примере рекомендаций CIS для Windows, которые включают сотни пунктов и требуют тщательного исполнения [5]. В таких условиях автоматизация аудита становится необходимой для обеспечения систематической и полной проверки безопасности. Национальный институт стандартов и технологий США (NIST) разработал методологию автоматизированного аудита, которая фактически стала стандартом де-факто для оценки соответствия информационных систем требованиям безопасности. В Российской Федерации проведение регулярного аудита безопасности информационных систем является одной из обязательных технических мер по обеспечению безопасности информационных систем, поэтому автоматизация этого процесса – актуальная задача [6-7].

После обзора литературы для исследования были выбраны SCAP (Security Content Automation Protocol) и OVAL (Open Vulnerability and Assessment Language), которые играют не последнюю роль в автоматизации аудита информационных систем, что позволяет эффективно управлять уязвимостями и конфигурациями [8-10]. SCAP

включает процессы поиска уязвимостей, исправления конфигураций и оценки безопасности. OVAL, как язык декларативных утверждений, описывает как уязвимости, так и необходимые состояния системы для обеспечения безопасности. Аудит с помощью OVAL профилей может быть адаптирован под различные задачи и системы, обеспечивая обнаружение уязвимостей, проверку конфигураций и соответствие требованиям стандартов безопасности. Для операционных систем, таких как ОС Альт, профили можно интегрировать через API репозитория, что значительно упрощает процесс автоматизированного контроля безопасности [11].

В рамках исследования был разработан скрипт, который реализует автоматизацию аудита информационной системы, собирая информацию об установленных пакетах, загружая OVAL профиль для проверки уязвимостей, выполняя сканирование и анализируя результаты. Он автоматически предоставляет рекомендации по обновлению уязвимых пакетов с использованием ALTRepo API для управления уязвимостями (Таблица 1).

Таблица 1 – Используемые API Endpoints

API Endpoint	Описание
GET /vuln/bdu	Возвращает информацию о текущих уязвимостях, связанных с базой данных уязвимостей (BDU), для пакетов.
GET /vuln/bdu/fixes	Возвращает список исправлений для уязвимостей BDU. Включает данные о версиях пакетов, для которых были выпущены исправления.
GET /vuln/cve	Предоставляет информацию о текущих уязвимостях, связанных с CVE (Common Vulnerabilities and Exposures), для конкретных пакетов.
GET /vuln/cve/fixes	Возвращает список исправлений для уязвимостей CVE. Включает данные о версиях пакетов, в которых эти уязвимости были исправлены.
GET /errata/export/oval/{branch}	Экспортирует OVAL профили для указанной ветки (branch) репозитория ALT Linux, которые используются для аудита уязвимостей и соответствия требованиям.

Однако при разработке возникли, связанных с некорректной структурой OVAL профиля (Таблица 2). Во-первых, атрибут «version» в элементах «definition» содержит некорректные значения, такие как oval:org.altlinux.errata:def:20191469, где требуется целое число, значения не соответствуют стандартам OVAL и вызывают ошибку. Во-вторых, элементы «textfilecontent54_test», «textfilecontent54_object» и «textfilecontent54_state» неожиданно появляются в контексте, где ожидаются другие элементы схемы, например, такие как «rpminfo_test». Эти элементы поддерживаются и активно используются в OVAL 5.11.1, но они должны быть правильно вписаны в схему, которая используется в профиле.

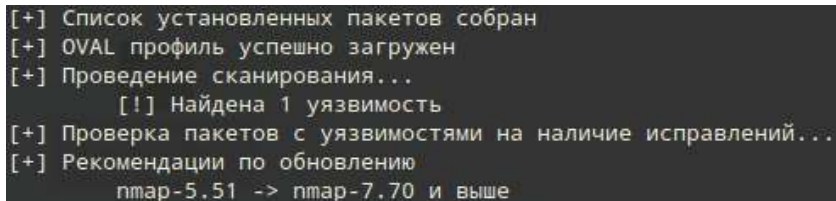
Таблица 2 – Проблемы со структурой OVAL профиля

Элемент	Проблема	Решение
<version>	Значение атрибута version содержит строковые значения (oval:org.altlinux.errata:def:20191469), что не соответствует требованиям схемы (должно быть числом).	Заменить строковые значения на целые числа, например, version="1".
<textfilecontent54_test>	Элемент не ожидается в текущем контексте схемы, нарушает структуру профиля.	Необходимо пересмотреть структуру XML файла.
<textfilecontent54_object>	Элемент не ожидается в текущем контексте схемы, нарушает структуру профиля.	Необходимо пересмотреть структуру XML файла.
<textfilecontent54_state>	Элемент не ожидается в текущем контексте схемы, нарушает структуру профиля.	Необходимо пересмотреть структуру XML файла.
Ссылки «test_ref» на объекты	Ссылки на тесты и объекты, такие как test_ref="oval:org.altlinux.errata:tst:2001", ссылаются на несуществующие тесты	Создать тесты под каждую ссылку в <tests> элементе <rpminfo_test>
Элемент <timestamp>	Элемент появляется в контексте, где его не должно быть (комментарий в журнале ошибок).	Необходимо соблюдать порядок элементов.

Проблемы также включают отсутствие ссылок на тесты или объекты в элементе «test_ref», что указывает на несоответствие ссылок внутри профиля. Профиль ссылается на несуществующие тесты или объекты, что вызывает ошибку «No match found for key-sequence».

Помимо описанного, в некоторых случаях элементы, например, «timestamp», появляются в местах, где их не должно быть в соответствии со схемой [12]. Структура XML в целом не соответствует схеме, что приводит к ошибкам при проверке схемы.

Несмотря на возникшие проблемы с корректностью OVAL профилей для ОС Альт, определенные результаты в рамках исследования были достигнуты. Используя самописный тестовый профиль, удалось провести базовый аудит безопасности, который хотя и не был полностью корректным – профиль нельзя было получить автоматически через ALTRepo API, однако предоставил важную информацию для дальнейших действий по обеспечению безопасности информационной системы. В рамках тестов была получена информация об установленных в системе пакетах и об их уязвимостях, а также выданы рекомендации об обновлении уязвимых версий пакетов (рисунок 1). Промежуточные результаты демонстрируют важность интеграции автоматизированных инструментов для управления уязвимостями в информационных системах.



```
[+] Список установленных пакетов собран
[+] OVAL профиль успешно загружен
[+] Проведение сканирования...
    [!] Найдена 1 уязвимость
[+] Проверка пакетов с уязвимостями на наличие исправлений...
[+] Рекомендации по обновлению
    ntpar-5.51 -> ntpar-7.70 и выше
```

Рис. 1 – Рекомендации по обновлению уязвимых версий пакетов

В заключение необходимо отметить, что автоматизация аудита информационных систем с использованием, например, протоколов SCAP и языка OVAL, является основополагающим для обеспечения безопасности современных информационных систем. Применение автоматизации позволяет эффективно управлять уязвимостями и обновлениями в системах, таких как ОС Альт, за счет интеграции с API, предоставляющими данные об уязвимостях и исправлениях. Хотя при реализации были выявлены проблемы, связанные с корректностью OVAL профилей, их устранение способствует более точной и своевременной оценке безопасности системы.

Литература:

1. ГОСТ Р. 53622-2009 «Информационные технологии (ИТ) [Электронный ресурс] // Информационно-вычислительные системы. Стадии и этапы жизненного цикла, виды и комплектность документов. – URL: https://allgosts.ru/03/120/gost_r_53622-2009. (дата обращения: 05.10.2024).
2. Уймин, А.Г. Цифровые двойники сетевых инфраструктур: точность, методы и практические решения // Радиотехнические и телекоммуникационные системы. – 2023. – № 3(51). – С. 44-52. – doi 10.24412/2221-2574-2023-3-44-52.
3. Файзулин Р.Ф., Демичев М.С., Маркевич И.В., Оголь А.Р., Бондарев А.С. Роль pci dss в обеспечении информационной безопасности [Электронный ресурс] // Актуальные проблемы авиации и космонавтики. – 2022. – С. 312-314. – URL: <https://cyberleninka.ru/article/n/rol-pci-dss-v-obespechenii-informatsionnoy-bezopasnosti>. (дата обращения: 05.10.2024).
4. Николаев А.С. Исследование подходов к автоматизации процесса моделирования угроз безопасности информации согласно методическим рекомендациям ФСТЭК России [Электронный ресурс] // Приоритетные направления инновационной деятельности в промышленности. – 2021. – С. 113-114. – URL: <https://elibrary.ru/item.asp?id=46462572/> (дата обращения: 05.10.2024).
5. Center for Internet Security (CIS) Benchmarks [Электронный ресурс]. – URL: <https://learn.microsoft.com/en-us/compliance/regulatory/offering-CIS-Benchmark/> (дата обращения: 05.10.2024).
6. Приказ ФСТЭК России от 18 февраля 2013 г. N 21 [Электронный ресурс]. – URL: <https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-18-fevralya-2013-g-n-21> (дата обращения: 08.10.2024).
7. Методический документ от 17 мая 2023 г. [Электронный ресурс]. – URL: <https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/metodicheskij-dokument-ot-17-maya-2023-g> (дата обращения: 08.10.2024).
8. D'Hondt A., Bahmad H. Understanding SCAP Through a Simple Use Case [Электронный ресурс] // Hakin9 IT Security Magazine. – 2016. – Т. 11. – С. 100-110. – URL: https://dial.uclouvain.be/memoire/ucl/en/object/thesis:8128/datastream/PDF_05/view (дата обращения: 07.10.2024).

9. Toussaint M., Krifa S., Panetto H. Industry 4.0 data security: A cybersecurity frameworks review [Электронный ресурс] //Journal of Industrial Information Integration. – 2024. – С. 100604. – URL: <https://www.sciencedirect.com/science/article/pii/S2452414X24000487> (дата обращения: 07.10.2024).

10. ALTRepo API [Электронный ресурс]. – URL: <https://rdb.altlinux.org/api> (дата обращения: 05.10.2024).

11. The OVAL® Language Specification [Электронный ресурс]. – URL: https://oval.mitre.org/language/version5.11/oval-language-specification_12-18-2014.pdf (дата обращения: 05.10.2024).

АНАЛИЗ ТЕМПЕРАТУРНОГО РЕЖИМА РАБОТЫ ОБОРУДОВАНИЯ В ТЕЛЕКОММУНИКАЦИОННОМ ШКАФУ С УЧЕТОМ ВНЕШНИХ КЛИМАТИЧЕСКИХ УСЛОВИЙ

П.Ю. Озолев

Новосибирский государственный технический университет,

г. Новосибирск, studio154@bk.ru

Научный руководитель: Достовалов Д.Н., к.т.н.

В статье рассматривается температурный режим работы оборудования, установленного в телекоммуникационном шкафу, с учетом внешних климатических условий. Разработана математическая модель, учитывающая внутренние тепловые потоки и внешние климатические факторы, такие как температура окружающей среды, влажность и солнечная радиация. Модель позволяет анализировать тепловую нагрузку и оптимизировать параметры системы охлаждения, а также формулировать рекомендации по размещению оборудования для предотвращения перегрева.

The article analyzes the thermal operating conditions of equipment installed in a telecommunications box, taking into account external climate factors. A mathematical model was developed, accounting for internal heat flows and external climate factors such as ambient temperature, humidity, and solar radiation. The model enables thermal load analysis, optimization of cooling system parameters, and the formulation of recommendations for equipment placement to prevent overheating.

Введение

Телекоммуникационное оборудование устанавливается в самых различных климатических и эксплуатационных условиях, от серверных комнат с контролируемым микроклиматом до уличных телекоммуникационных шкафов в условиях высокой влажности, сильных перепадов температур и воздействием прямого солнечного

излучения. Поддержание оптимального температурного режима оборудования становится важной задачей, поскольку перегрев или недостаточное охлаждение может значительно снизить эффективность и надежность компонентов, а также повысить энергозатраты, связанные с их охлаждением.

Постановка задачи

Температурный режим телекоммуникационного оборудования определяется внутренними и внешними факторами. Внутренние факторы включают тепловыделение оборудования в зависимости от его рабочей нагрузки, плотность размещения устройств, а также наличие и мощность системы охлаждения. Внешние факторы связаны с климатическими условиями: температурой окружающей среды, уровнем солнечного излучения, влажностью и направлением потоков воздуха.

Одной из особенностей телекоммуникационных шкафов является то, что оборудование в них работает круглосуточно и находится в замкнутом пространстве, где тепло скапливается, что усугубляет проблему. При этом оборудование, размещенное в верхней части стойки, как правило, подвергается большей температурной нагрузке из-за конвекции и ограничения воздушного потока. Данные особенности делают необходимым создание эффективной модели теплового анализа, учитывающей различные условия эксплуатации и позволяющей оптимизировать параметры системы охлаждения.

Целью данного исследования является разработка и анализ математической модели, которая учитывает как внутренние, так и внешние климатические факторы, влияющие на температурный режим внутри шкафа. С помощью численных расчетов модель позволяет оценить тепловую нагрузку при различных климатических условиях, спрогнозировать поведение системы охлаждения и выработать рекомендации для размещения оборудования.

При решении задачи необходимо учесть следующие факторы:

1. Внутренние тепловые потоки, возникающие при работе оборудования.
2. Внешние климатические условия (температура, влажность, солнечное излучение и др.) которые варьируются в зависимости от времени года, времени суток и географического положения объекта.
3. Тепловой баланс, включающий расчет распределения температуры с учетом конвективных и радиационных теплопотоков. Модель должна учитывать неравномерное распределение температуры, возникающее из-за особенностей конструкции шкафа и положения оборудования.

Методология

Для моделирования теплового режима использованы методы теплопередачи, включающие кондукцию, конвекцию и излучение. Основные уравнения теплопередачи для анализа температурного режима:

1. Уравнение теплопередачи через кондукцию (1).

$$q = -k \frac{\Delta T}{\Delta x}, \quad (1)$$

где q – поток тепла через материал ($\text{Вт}/\text{м}^2$), k – коэффициент теплопроводности материала шкафа ($\text{Вт}/(\text{м} \cdot \text{К})$), ΔT – разность температур по разные стороны материала шкафа (К), Δx – толщина материала (м).

2. Уравнение теплового баланса (2).

$$C \frac{dT}{dt} = P_{\text{вых}} - hA(T - T_{\text{окр}}), \quad (2)$$

где C – теплоемкость оборудования ($\text{Дж}/\text{К}$), T – температура оборудования (К), t – время (с), $P_{\text{вых}}$ – мощность тепловыделения оборудования (Вт), h – коэффициент теплоотдачи ($\text{Вт}/(\text{м}^2 \cdot \text{К})$), A – площадь поверхности оборудования (м^2), $T_{\text{окр}}$ – температура окружающей среды (К).

3. Тепловая модель с учетом внешнего излучения. Солнечное излучение, нагревающее шкаф, может быть учтено добавлением радиационной составляющей (3).

$$Q_{\text{солн}} = \alpha I_{\text{солн}} A_{\text{внеш}} \cos(\theta), \quad (3)$$

где $Q_{\text{солн}}$ – мощность солнечного излучения, поглощаемого поверхностью шкафа (Вт), α – коэффициент поглощения материала шкафа, $I_{\text{солн}}$ – интенсивность солнечного излучения ($\text{Вт}/\text{м}^2$), $A_{\text{внеш}}$ – площадь внешней поверхности шкафа, подверженной излучению (м^2), θ – угол падения солнечных лучей на поверхность шкафа.

Экспериментальная часть

Для оценки теплового режима телекоммуникационного оборудования было проведено тестирование в различных условиях внешней температуры и интенсивности тепловыделения. В таблице 1 представлены результаты тестирования, включая значения внутренней температуры в шкафу и мощность охлаждения. Эти данные позволяют оценить влияние климатических факторов на температурный режим и определить оптимальные параметры для системы охлаждения.

Таблица 1 – Температурные условия работы оборудования

Внешняя температура (°C)	Температура внутри стойки (°C)	Тепловыделение оборудования (Вт)	Мощность охлаждения (Вт)
10	25	200	50
20	35	400	100
35	50	600	200
45	60	800	300

Результаты и обсуждение

Рассмотрение влияния внешней температуры, влажности и солнечной радиации позволило выявить, что температура внутри шкафа значительно возрастает при повышении температуры окружающей среды. Наибольшая разница температур была отмечена в дневное время и в зонах, подверженных солнечному излучению.

Выполненный численный анализ с использованием вышеуказанных формул показал, что увеличение мощности системы охлаждения становится критическим при внешних температурах выше 30°C и высокой солнечной радиации. Важно отметить, что при моделировании температурных изменений в зависимости от суточных и сезонных колебаний окружающей температуры можно определить оптимальные параметры для систем охлаждения.

На основе разработанной модели теплового анализа были сформулированы рекомендации по выбору системы охлаждения и размещению оборудования. Полученные данные позволяют:

1. Предотвратить перегрев оборудования путем оптимального размещения устройств, генерирующих наибольшую тепловую нагрузку, ближе к источникам охлаждения.

2. Снизить энергозатраты на охлаждение за счет рационального использования вентиляторов и дополнительного оборудования для поддержания стабильного температурного режима, что особенно актуально для регионов с высокими среднегодовыми температурами.

3. Продлить срок службы компонентов путем минимизации термических стрессов, вызванных колебаниями температуры, и улучшения условий их эксплуатации.

Внедрение подобной тепловой модели в процесс проектирования и эксплуатации телекоммуникационных систем позволяет заранее оценить возможные риски и оптимизировать ресурсы. Дальнейшие исследования в этом направлении могут быть направлены на разработку систем автоматического мониторинга и управления температурой на основе IoT-датчиков и алгоритмов машинного обучения. Это позволит в

реальном времени адаптировать систему охлаждения к текущим климатическим условиям и нагрузкам, обеспечивая непрерывность и эффективность работы оборудования.

Литература:

1. Иващенко А.А., Романова М.Л. Температурные режимы работы телекоммуникационного оборудования: практическое руководство. – М.: Энергия, 2015. – 320 с.
2. Левченко В.С., Петров А.И. Теплообмен и вентиляция в телекоммуникационных системах. – СПб.: ТехноПресс, 2019. – 280 с.
3. Иванов П.А., Сидорова Л.К. Анализ температурного режима телекоммуникационного оборудования в стойках // Журнал технической физики. – 2020. – № 4. – С. 55–63.
4. Козлов А.В., Шишкин Н.И. Оценка эффективности систем охлаждения в условиях повышенных температур // Вестник инженерных наук. – 2018. – № 3. – С. 112–119.
5. Соколова Н.П., Климов В.Ф. Влияние климатических условий на надежность оборудования // Электронная техника и автоматика. – 2016. – № 1. – С. 28–34.

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ СИСТЕМА ПОДДЕРЖКИ УЧЕТА НА ПРЕДПРИЯТИЯХ СЕЛЬСКОХОЗЯЙСТВЕННОЙ ОТРАСЛИ

Е.В. Яковлев

**Донской государственный технический университет,
г. Ростов-на-Дону, yakovlev.evgeny09@gmail.com**

Научный руководитель: Климова Д.Н., к.т.н., доцент

В условиях динамичного развития сельского хозяйства и усложнения требований к обработке данных такие системы становятся ключевым инструментом для оптимизации учета ресурсов, планирования, мониторинга, анализа и принятия управленческих решений. В работе рассматриваются функциональные возможности информационно-аналитических систем (ИАС), их архитектура, методологические подходы к интеграции с основными бизнес-процессами агропредприятия, а также аспекты их адаптации к специфике аграрного сектора. Основное внимание уделяется вопросам автоматизации учета производственных показателей, оперативного анализа, что позволяет агропредприятиям повышать производительность, снижать затраты на агрохимию и улучшать качество принимаемых решений.

In the context of the dynamic growth of agriculture and increasing requirements for data processing, such systems become essential tools for optimizing resource

accounting, planning, monitoring, analysis, and decision-making processes. This work examines the functional capabilities of information-analytical systems (IAS), their architecture, methodological approaches to integration with core business processes of agricultural enterprises, and aspects of adaptation to the agricultural sector's specific needs. The emphasis is on automating the accounting of production indicators, enabling prompt analysis and forecasting, which allows agricultural enterprises to increase productivity, reduce costs, and improve the quality of decision-making.

Введение

Одна из важнейших сфер в современном мире — это сельское хозяйство, которое играет ключевую роль в обеспечении продовольственной безопасности и устойчивого развития мировой экономики. Для эффективного управления сельскохозяйственными предприятиями и повышения их производительности необходимо активно внедрять современные информационные технологии.

Важным фактором успешной деятельности в современном сельском хозяйстве является контроль и учет использования химических ресурсов, включая удобрения и пестициды. Сложность и объем учета химических ресурсов растут, требуя специализированных информационных решений для эффективного управления, поэтому разработка средств автоматизации учетно-аналитической деятельности по использованию химической продукции является актуальной задачей.

Цель данной работы — создание системы автоматизации ведения информационно-аналитического учета агро-химических работ на сельскохозяйственных предприятиях.

В качестве среды реализации выбрана платформа 1С:Предприятие 8, которая предлагает гибкие и мощные средства разработки информационно-аналитических систем. С ее помощью можно оптимизировать учет химической продукции и обеспечить оптимальное использование ресурсов на сельскохозяйственных предприятиях.

Учет химических продуктов в сельскохозяйственном предприятии.

Сельское хозяйство — это не просто производство продуктов питания. Это сложная отрасль, где каждый фактор играет решающую роль в успехе урожая. Вот почему так важно эффективно управлять химическими ресурсами, такими как удобрения, пестициды и другие агрохимикаты.

Именно поэтому было решено заняться этой темой. Моя работа направлена на создание современной информационно-аналитической системы, которая поможет эффективно управлять химическими ресурсами на ферме или хозяйстве. Необходимо, чтобы фермеры и

хозяева предприятий могли легко отслеживать, сколько ресурсов используется, какие остатки остаются, и как можно улучшить эффективность использования этих ресурсов.

Важно понимать, что такая система не только поможет сэкономить время и ресурсы, но и способствует более эффективному использованию земли и защите окружающей среды [1]. Это позволит сельскохозяйственным предприятиям быть более конкурентоспособными и устойчивыми в условиях современного рынка.

Индивидуальные особенности:

1) инновационный подход. Проведенное исследование представляет собой инновационный подход к учету химии на сельскохозяйственных предприятиях. Разработка информационно-аналитической системы на платформе 1С позволяет создать современное и эффективное решение для учета химических ресурсов;

2) гибкость и адаптивность. Разработанная конфигурация на платформе 1С обладает гибкостью и адаптивностью, что позволяет ее легко настраивать под конкретные потребности каждого сельскохозяйственного предприятия;

3) интеграция и совместимость. Система разработана с учетом возможности интеграции с другими информационными системами и управленческими решениями, что обеспечивает ее совместимость с существующими инфраструктурами предприятий.

Значимость автоматизации учетной деятельности.

Значимость автоматизации учетной деятельности на сельскохозяйственных предприятиях трудно переоценить. Вот несколько ключевых аспектов:

1) точность и надежность данных. Автоматизация учета позволяет исключить ошибки, связанные с ручным вводом и обработкой информации. Это гарантирует точность и надежность данных, что является критически важным для принятия правильных управленческих решений [2];

2) экономия времени и ресурсов. Автоматизация учета существенно сокращает время, затрачиваемое на рутинные операции по записи, отслеживанию и анализу данных;

3) увеличение эффективности процессов. Автоматизация учета позволяет оптимизировать бизнес-процессы, ускорить выполнение задач;

4) лучшее управление ресурсами. Благодаря автоматизированной учетной системе управленцы получают более полную и своевременную

информацию о состоянии ресурсов, что позволяет более эффективно планировать и распределять их использование.

Функциональные требования

Функциональные требования определяют, каким образом система должна взаимодействовать с пользователями и выполнять свои функции. В случае информационно-аналитической системы поддержки учета химии на сельскохозяйственном предприятии в 1С, функциональные требования включают в себя следующее:

- 1) учет и контроль агрохимикатов:
 - а) ведение учета агрохимикатов, включая приобретение, использование и остатки;
 - б) мониторинг сроков годности агрохимикатов и предупреждение о превышении срока годности;
 - в) отслеживание расходов агрохимикатов по сельскохозяйственным участкам и культурам.
- 2) управление складскими запасами:
 - а) учет остатков агрохимикатов на складах;
 - б) оперативное обновление данных об остатках при приобретении и использовании агрохимикатов;
 - в) генерация уведомлений о необходимости пополнения запасов агрохимикатов.
- 3) планирование и анализ использования химии:
 - а) анализ эффективности использования агрохимикатов на основе данных учета;
 - б) формирование отчетов о расходе химии по периодам, участкам и видам культур;
 - в) Предоставление рекомендаций по оптимизации использования агрохимикатов для повышения эффективности производства.
- 4) интеграция с другими системами:
 - а) обеспечение интеграции с системами учета земельных участков, культур, оборудования и другими информационными системами предприятия;
 - б) передача данных о расходе агрохимикатов для бухгалтерского учета и анализа финансовой эффективности.

Выводы

В заключение, можно отметить, что разработка и внедрение информационно-аналитической системы учета на предприятиях сельскохозяйственной отрасли предоставляет широкие возможности для оптимизации управленческих и учетных процессов. Такие системы

позволяют автоматизировать сбор и обработку данных, что существенно сокращает затраты времени и снижает вероятность ошибок, связанных с человеческим фактором. Информационно-аналитические системы способствуют повышению эффективности управления ресурсами, улучшению планирования и прогнозирования, а также повышению общей производительности и устойчивости агропредприятий. В условиях растущей конкуренции и усложняющихся требований рынка такие технологии становятся важным инструментом для обеспечения стабильного роста и достижения стратегических целей в сельском хозяйстве.

Литература:

1. *Гончаров Д.И., Хрусталева Е.Ю.* Технологии интеграции «1С: Предприятие 8.2». – 1С-Паблишинг, 2011. – 358 с.
2. Обзор компьютерных систем бизнес-аналитики 2016 г. [Электронный ресурс]. – URL: <http://www.e-postulat.ru/index.php/Postulat/article/view/50/53> (дата обращения: 01.05.2024).
3. 1С: Предприятие 8.3: комьюнити-лицензия [для разработки и отладки прикладных решений (конфигураций) на базе технологической платформы 1С:Предприятие] / разработчик «1С». – Москва. – Электронная программа: электронная.
4. *Зыков, С.В.* Программирование: учебник и практикум для вузов / С.В. Зыков. – Москва: Издательство Юрайт, 2021. – 320 с. – (Высшее образование).
5. *Бартенев, О.В.* 1С: Предприятие: программирование для всех / О.В. Бартенев. - М.: Диалог-Мифи, 2016. - 464 с.
6. *Радченко, М.Г.* 1С: Предприятие 8.3. Практическое пособие разработчика. Примеры и типовые приемы / М.Г. Радченко, Е.Ю. Хрусталева. – М.: 1С-Паблишинг, 2011. - 965 с.
7. *Митичкин, С.А.* Разработка в системе 1С: Предприятие 8.3 - СПб.: «1С». 2017-384 с.
8. *Хрусталева Е.Ю., Радченко М.Г.* Язык запросов "1С: Предприятия 8" / Е.Ю. Хрусталева, Радченко М. Г.: 000 "1 С-Паблишинг", 2013. – 358 с.
9. *Гончаров Д.И., Хрусталева Е.Ю.* Технологии интеграции «1С: Предприятие 8.2». – 1С-Паблишинг, 2011. – 358 с.
10. *Кашиев С.М.* Программирование в 1С: Предприятие 8.3. – СПб.: Питер, 2014. – 304 с.

НАУЧНОЕ НАПРАВЛЕНИЕ – ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ОБРАБОТКИ ДАННЫХ

Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, АНАЛИЗ И ОБРАБОТКА ДАННЫХ

ПРИМЕНЕНИЕ ВЕКТОРНЫХ БАЗ ДАННЫХ ДЛЯ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ НА ПРИМЕРЕ ИСПОЛЬЗОВАНИЯ БАЗЫ QDRANT

К.А. Шнайдер

Новосибирский государственный технический университет,
г. Новосибирск, k.shnajder@corp.nstu.ru

Научный руководитель: Стасышин В.М., к.т.н., доцент

В условиях стремительного увеличения объемов и сложности обрабатываемых данных современные базы данных сталкиваются с необходимостью поиска и анализа информации на основе семантики, а не только совпадений. Это послужило толчком к развитию и использованию векторных баз данных (ВБД), которые широко применяются в сервисах для маркетинга, медицины, искусственного интеллекта. С помощью векторных баз осуществляется более быстрый и точный поиск схожих объектов. В данной работе рассматриваются преимущества ВБД QDrant – одного из лидирующих решений по скорости внедрения, и его применение в задаче семантического поиска текстов.

With the rapid increase in the volume and complexity of processed data, modern databases are faced with the need to search and analyze information based on semantics, not just matches. This has spurred the development and use of vector databases (VDBs), which are widely used in services for marketing, medicine, and artificial intelligence. Vector databases are used to search for similar objects more quickly and accurately. This paper discusses the advantages of QDrant VDB, one of the leading solutions in terms of speed of implementation, and its application in the task of semantic text search.

Современные информационные системы требуют анализа больших объемов информации, а также поиска на основе семантического сходства – людям чаще всего важны не точные совпадения данных, а их контекстное содержание. Классический подход при работе с данными не всегда удовлетворяет этому требованию, поэтому появились и активно развиваются векторные базы данных. Такие базы позволяют хранить данные в виде векторов (эмбедингов), то есть многомерных представлениях, отражающих информацию о контексте. Векторные

модели позволяют на основании метрик искать не просто эквивалентные, но и близкие по смыслу данные – что довольно востребовано для построения рекомендаций, анализа изображений, видео и звуков.

I. Определение векторной базы данных и её отличие от реляционной.

Векторные базы данных, в отличие от реляционных, создают и хранят данные в форме векторов – многомерных точек в пространстве, где каждая из точек несет определённое смысловое значение. Такой подход позволяет использовать метод поиска ближайших соседей для оценки схожести между объектами, на базе которого можно строить семантический поиск. Реляционные базы предназначены для точного поиска и установления четких связей, тогда как ВБД опираются на выявление неявных связей и поиска близости между объектами. К числу наиболее популярных векторных баз относятся следующие: QDrant, Pinecone, Weaviate и Milvus. Qdrant выделяется среди остальных за счет открытого исходного кода и высокопроизводительных алгоритмов, что делает его удобным для интеграции и настройки.

II. QDrant: назначение, особенности и внутренние механизмы.

Как уже упоминалось, QDrant – это ВБД с открытым исходным кодом, специально разработанная для высокопроизводительного поиска по схожести между векторами. Основной механизм QDrant заключается в эффективной индексации векторов и поддержке масштабируемости.

QDrant поддерживает гибридные модели данных, позволяя работать как с векторными, так и с традиционными атрибутами. Ключевыми преимуществами QDrant являются простота интеграции с различными приложениями, высокая точность поиска и поддержка современных методов машинного обучения, что позволяет гибко настраивать и адаптировать базу под конкретные прикладные задачи.

Внутренний механизм поиска в QDrant содержит несколько этапов:

1. Индексирование векторов: одной из популярных структур для индексации является HNSW (Hierarchical Navigable Small World) – она основана на графе, который позволяет быстро искать векторы, близкие к запрашиваемому, с минимальными затратами. В процессе построения этого индекса QDrant создает уровни, по которым иерархически распределяются векторы, поиск по этим уровням идёт от более обобщенных векторов (верхний уровень) до наиболее детализированных (нижний уровень).

2. Поиск ближайших соседей: QDrant использует индекс HNSW для нахождения ближайших соседей, ориентируясь на метрику расстояния, такую как косинусное расстояние или евклидова мера (можно использовать и другие метрики).

3. Оптимизация поиска: QDrant поддерживает многопоточную обработку запросов, а также имеет систему кэширования, которая позволяет хранить результаты наиболее частых запросов. Также кроме описанного индекса HNSW можно использовать другие алгоритмы и методы настройки для различных типов данных и задач.

III. Пример решения задачи семантического поиска с использованием QDrant.

Рассмотрим практическое использование QDrant для решения задачи семантического поиска по научным и проектным работам. В качестве входных данных поисковая система получает текстовое наименование работы, в качестве результата поиска пользователь получает перечень похожих работ и их авторов. Основной целью задачи является поиск авторов релевантных по смыслу материалов, которые работали или работают в искомой предметной области. Наименования научных и проектных работ могут использовать различные формулировки и термины, поэтому классический поиск, основанный на сопоставлении и сравнении, не подходит для решения задачи.

Для решения задачи был реализован следующий процесс (Рисунок 1).

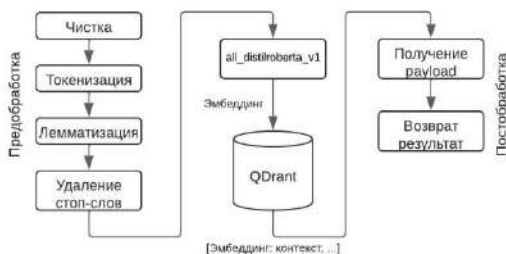


Рис. 1. – процесс получение результата на запрос пользователя в поисковой системе

1. Предобработка текста: на первом этапе происходит очистка и подготовка отправленного пользователем текста к преобразованию в векторное представление. Процесс включает: очистку текста от лишних символов, токенизацию, лемматизацию, удаление стоп-слов (для русского и английского языка), создание эмбедингов с помощью предобученной модели all-distilroberta-v1, построенной на основе DistilRoBERTa, которая является модификацией BERT.

2. Поиск в QDrant: после преобразования текста в вектор (эмбединг) выполняется поиск в базе данных с использованием косинусного расстояния в качестве метрики. Процесс поиска находит

векторы, которые наиболее близки по направлению к запросу вектору, позволяя получить наиболее релевантные результаты.

3. Получение результатов и доступ контексту: система возвращает только те векторы, которые имеют наибольшее значение косинусного расстояния. Дальнейшая работа происходит с наиболее релевантными векторами, где QDrant позволяет обратиться к связанным с ними исходным данным (название и автор работы).

4. Формирование и возврат результатов: на последнем этапе формируется итоговый отправляемый список работ и их авторов, что позволяет пользователю найти не только схожие работы, но и увидеть, кто из авторов работает в одной и той же предметной области.

Интерфейс разработанной системы поиска представлен ниже (Рисунок 2).

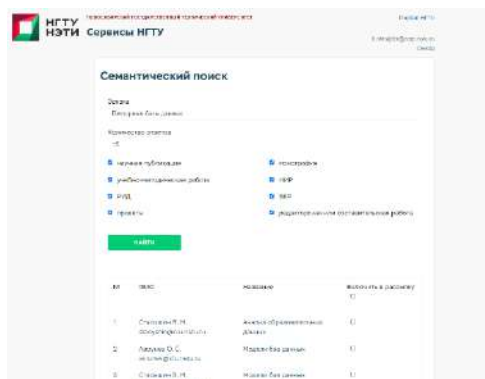


Рис. 2. – интерфейс разработанной поисковой системы

Таким образом, описанный подход построения семантического поиска с использованием QDrant может быть использован в смежных задачах. Например, подбор кандидатов для проведения собеседований на основании их соответствия стеку технологий и компетенций, поиск похожих фильмов или книг на основании их описания, создание play-листа на базе прослушанных пользователем композиций, поиск в магазинах одежды похожих товаров на основании содержимого корзины и многое другое.

Литература:

1. *Аврунев О.Е., Стасышин В.М.* Модели баз данных: [учебное пособие] [Электронный ресурс] / Новосибирский государственный технический университет. – Новосибирск: Изд-во НГТУ, 2018. – 121, [2]

с.: ил. – URL: http://elibrary.nstu.ru/source?bib_id=vtls000239752 (дата обращения: 27.10.2024).

2. Басинов А.А., Демич О.В. Семантический поиск: проблемы и технологии [Электронный ресурс] // Вестник АГТУ. Серия: Управление, вычислительная техника и информатика. – 2012. – №1. – URL: <https://cyberleninka.ru/article/n/semanticheskii-poisk-problemy-i-tehnologii> (дата обращения: 27.10.2024).

3. QDrant [Электронный ресурс] // QDrant. – URL: <https://qdrant.tech> (дата обращения: 27.10.2024).

ИЗМЕРЕНИЕ ПРОЗРАЧНОСТИ СЛАБО ЗАМУТНЕННОЙ АТМОСФЕРЫ С ПРИМЕНЕНИЕМ ЛИДАРНЫХ ТЕХНОЛОГИЙ

Д.Д. Ковригин, И.А. Потапова, Т.Ю. Яковлева
Российский государственный
гидрометеорологический университет,
г. Санкт-Петербург, Россия, potapovaira@yandex.ru

Статья посвящена анализу методов интерпретации лидарной информации в случаях слабо замутненной атмосферы. При таких условиях измерение сигналов обратного рассеяния малой мощности становится сложным из-за значительного влияния фоновой засветки. Основной целью исследования является развитие и усовершенствование методов обработки эхо-сигналов. В статье представлен метод обработки лидарных данных, который не требует итерационных процессов.

The article is devoted to the analysis of methods of interpretation of lidar information in cases of slightly cloudy atmosphere. Under such conditions, measuring low-power backscattering signals becomes difficult due to the significant influence of background illumination. The main purpose of the research is the development and improvement of echo signal processing methods. The article presents a lidar data processing method that does not require iterative processes.

Введение

В данной работе рассматривается проблема интерпретации лидарной информации при зондировании слабо замутненной атмосферы, где фоновая засветка значительно влияет на измерения сигналов обратного рассеяния малой мощности.

Целью исследования является развитие методов обработки эхо-сигналов, необходимых для эффективного анализа атмосферных условий с повышенной прозрачностью воздуха. В предыдущих работах, таких как [1-4], использовался метод последовательных приближений для обработки лидарных данных, что ограничивало оперативность

анализа. В данной работе предлагается новый метод, который устраняет этот недостаток, обеспечивая более быструю и точную интерпретацию лидарных сигналов. Результаты исследования могут способствовать улучшению качества данных, получаемых с помощью лидарных систем, и расширению их применения в метеорологии и экологии.

Обработка сигналов обратного рассеяния с помощью итераций

Обработка данных, полученных с помощью лидаров, осуществляется на основе решения лидарного уравнения, которое устанавливает связь между сигналом обратного рассеяния и коэффициентами ослабления и обратного рассеяния. В условиях слабо замутненной оптически однородной атмосферы, когда коэффициенты ослабления и обратного рассеяния остаются постоянными вдоль зондируемой трассы [5-8], лидарное уравнение можно представить в следующем виде:

$$P_i = P_* + \frac{B}{R_i^2} \exp(-2\sigma R_i), \quad (1)$$

где

$$B = A\beta. \quad (2)$$

A – постоянная лидара, β – коэффициент обратного рассеяния, σ – коэффициент ослабления. P – мощность сигнала обратного рассеяния, P_* – мощность солнечного излучения, рассеянного атмосферой в направлении на приемное устройство лидара, R – расстояние между лидаром и i – М рассеивающим элементом.

Задача заключается в определении неизвестных постоянных P_* , B , σ . Чтобы определить их значения с наименьшей погрешностью, нужно применить метод наименьших квадратов. В этом случае необходимо минимизировать сумму:

$$\delta^2 = \sum_1^n \left(P_* + \frac{B}{R_i^2} \exp(-2\sigma R_i) - P_i \right)^2, \quad (3)$$

а это подразумевает использование, в том числе, итерационного процесса из-за нелинейного характера задачи.

Обработка сигналов обратного рассеяния без итерационного процесса

Для предлагаемого решения, вместе с уравнением (1), необходимо рассмотреть уравнения

$$P_{i+1} = P_* + \frac{B}{R_{i+1}^2} \exp(-2\sigma R_{i+1}), \quad (4)$$

$$P_{i+2} = P_* + \frac{B}{R_{i+2}^2} \exp(-2\sigma R_{i+2}). \quad (5)$$

С учетом уравнений (1), (4), (5) можно получить:

$$\frac{(P_i - P_*)R_i^2}{(P_{i+1} - P_*)R_{i+1}^2} = \frac{(P_{i+1} - P_*)R_{i+1}^2}{(P_{i+2} - P_*)R_{i+2}^2} = \exp(2\sigma h), \quad (6)$$

где h – шаг,

$$h = R_{i+1} - R_i. \quad (7)$$

С учетом уравнений (6), вместо суммы (1) можно минимизировать сумму

$$\delta_*^2 = \sum_1^n a_i P_*^2 + b_i P_* + c_i, \quad (8)$$

где

$$a_i = R_{i+1}^4 - R_i^2 R_{i+2}^2, \quad (9)$$

$$b_i = (P_i + P_{i+2})R_i^2 R_{i+2}^2 - 2P_{i+1} R_{i+1}^4, \quad (10)$$

$$c_i = P_{i+1}^2 R_{i+1}^4 - P_i P_{i+2} R_i^2 R_{i+2}^2. \quad (11)$$

Для расчета фоновой засветки, с учетом полученных выражений, получится следующее:

$$P_*^3 + aP_*^2 + bP_* + c = 0, \quad (12)$$

где

$$a = \frac{3 \sum_1^n a_i b_i}{2 \sum_1^n a_i^2}, \quad b = \frac{\sum_1^n (2c_i + a_i) a_i}{2 \sum_1^n a_i^2}, \quad c = \frac{\sum_1^n b_i c_i}{2 \sum_1^n a_i^2}. \quad (13)$$

Решение Кардано уравнения (12):

$$P_* = a_* + b_* - \frac{a}{3}, \quad (14)$$

где

$$a_* = \sqrt[3]{-\frac{q}{2} + \sqrt{Q}}, \quad b_* = \sqrt[3]{-\frac{q}{2} - \sqrt{Q}}, \quad (15)$$

$$Q = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2, \quad p = -\frac{a^2}{3} + b, \quad q = 2\left(\frac{a}{3}\right)^3 - \frac{ab}{3} + c. \quad (16)$$

Результаты определения фоновой засветки, в качестве примера для $R_0 = 2,5 \text{ км}$, представлены в таблице 1.

**Таблица 1 – Результаты определения фоновой засветки
для $R_0 = 2,5 \text{ км}$**

$\Delta R, \text{ км}$	1	4	6	8
$P_*/10 \text{ шум}$	38,06	37,65	37,15	37,12

Результаты определения фоновой засветки для $R_0 = 10,5 \text{ км}$ представлены в таблице 2.

**Таблица 2 – Результаты определения фоновой засветки
для $R_0 = 10,5 \text{ км}$**

$\Delta R, \text{ км}$	0,5	1	2	2,5
$P_*/10 \text{ шум}$	36,97	36,99	36,99	36,99

Следует отметить наличие систематической погрешности для $R_0 = 2,5 \text{ км}$ и ее отсутствие для $R_0 = 10,5 \text{ км}$ [9-12].

Заключение

В результате проведенного исследования было выявлено, что интерпретация лидарной информации в условиях слабо замутненной атмосферы представляет собой значительную проблему, особенно из-за влияния фоновой засветки на сигналы обратного рассеяния малой мощности. Цель работы заключалась в разработке новых методов обработки эхо-сигналов, что является важным шагом для повышения точности и оперативности анализа атмосферных данных. Предложенный метод, свободный от недостатков, связанных с использованием последовательных приближений, позволяет значительно улучшить скорость обработки лидарных данных. Это открывает новые возможности для применения лидарных технологий в метеорологии и экологии, обеспечивая более надежные результаты в условиях повышенной прозрачности воздуха. В целом, результаты работы способствуют более глубокому пониманию процессов, происходящих в атмосфере, и могут быть полезны для дальнейших исследований в данной области.

Литература:

1. Yegorov A.D., Potapova I.A., Rzhonsnitskaya Yu.B., Sanotskaya N.A., Shchadin A.V. Atmospheric aerosol measurements and reliability problem: new results // International Journal of Remote Sensing. – 2014. – 35. – P. 5750-5765.

2. Егоров А.Д., Потапова И.А. Лидарные исследования прозрачности атмосферы // Труды Главной геофизической обсерватории им. А.И. Воейкова. – № 553. – 2004. – С. 131-142.
3. Егоров А.Д., Потапова И.А. Способ определения прозрачности атмосферы // Патент на изобретение RU 2439626. – 2012. – С2, 10.01.2012. Заявка № 2009144060/28 от 27.11.2009.
4. Егоров А.Д., Потапова И.А., Ржонсницкая Ю.Б., Драбенко В.А. Лидарное зондирование атмосферного аэрозоля // Метеорологический вестник. – Т. 5. – № 2. – 2013. – С. 42-51.
5. Егоров А.Д., Потапова И.А., Ржонсницкая Ю.Б. Обращение лидарных сигналов малой мощности // Оптический журнал. – 2007. – Т. 74, № 10. – С. 25-28.
6. Egorov A.D., Potapova I.A., Shchukin G.G. Lidar methods for probing an atmospheric aerosol // Journal of Optical Technology. – 2001. – Т. 68, № 11. – С. 801-804.
7. Дьяченко Н.В., Егоров А.Д., Потапова И.А., Скобликова А.Л., Саноцкая Н.А., Яковлева Т.Ю. Интерпретация слабого радиационного излучения при лидарном зондировании атмосферы // Сборник тезисов международного симпозиума «Атмосферная радиация и динамика» (МСАРД-2019) СПб, 2019. – С. 189-190.
8. Бобровский А.П., Дьяченко Н.В., Егоров А.Д., Косцов В.В., Калиничев Д.В., Михтеева Е.Ю., Потапова И.А., Скобликова А.Л., Хлябич П.П., Яковлева Т.Ю. Влияние фоновой засветки на точность определения прозрачности слабо замутненной атмосферы // Информационные технологии и системы: управление, экономика, транспорт, право. – № 4 (36). – 2019. – С. 49-52.
9. Бобровский А.П., Дьяченко Н.В., Косцов В.В., Михтеева Е.Ю., Потапова И.А., Скобликова А.Л., Хлябич П.П., Яковлева Т.Ю. Лидарная прозрачнострия слабо замутненной атмосферы с учетом фоновой засветки // Информационные технологии и системы: управление, экономика, транспорт, право. № S1, 2022. – С. 47-50.
10. Потапова И.А., Волкова Н.А., Михтеева Е.Ю., Потапова В.Р., Хлябич П.П. Анализ воздействия естественного излучения атмосферы на принимаемый сигнал при исследовании прозрачности // В сборнике: Инновационные методы математики и физики в экологических и гидрометеорологических исследованиях. Сборник трудов Международной научно-практической конференции. Санкт-Петербург, 2024. – С. 160-167.
11. Бобровский А.П., Дьяченко Н.В., Михтеева Е.Ю., Потапова И.А., Потапова В.Р., Скобликова А.Л., Хлябич П.П., Яковлева Т.Ю. Влияние воздействия естественного рассеянного излучения атмосферы на отраженный сигнал при исследовании атмосферы низкой мутности // В

сборнике: Гидрометеорология и физика атмосферы: современные достижения и тенденции развития. Материалы II Международной научно-практической конференции. Санкт-Петербург, 2024. С. 34-39.

12. Махлин Н.К., Потапова И.А., Ягнюков С.А. Обратная задача лидарного зондирования слабо замутненной атмосферы // В сборнике: Инновационные методы математики и физики в экологических и гидрометеорологических исследованиях. Сборник трудов Всероссийской научно-практической конференции. Под редакцией И.В. Зайцевой. Санкт-Петербург, 2023. С. 129-137.

АНАЛИЗ НАДЕЖНОСТИ ТАБЛЕТПРЕССА GZP-32 НА ОСНОВЕ ВИНЕРОВСКОЙ ДЕГРАДАЦИОННОЙ МОДЕЛИ

Е.А. Осинцева

**Новосибирский государственный технический университет,
г. Новосибирск, osinceva.j@gmail.com**

Научный руководитель: Чимитова Е.В., д.т.н., доцент

В данной статье рассмотрена задача анализа надежности таблетпресса GZP-32. Построена винеровская деградационная модель с ковариатами, получен оптимальный план эксперимента и сформулированы рекомендации для повышения отказоустойчивости оборудования.

The problem of reliability analysis of the GZP-32 Tablet Press has been considered in this paper. The Wiener degradation model with covariates has been constructed, the optimal experimental design has been obtained, and recommendations for increasing the lifetime of the equipment were formulated.

Современные реалии мировой экономики, связанные с введением санкций против Российской Федерации, привели к значительным изменениям в цепочках поставок и логистике. Одним из наиболее ощутимых последствий стало усложнение ввоза комплектующих для промышленного оборудования, что оказывает непосредственное влияние на стабильность производственных процессов. В условиях дефицита запчастей и компонентов предприятия вынуждены искать новые подходы для минимизации простоев и повышения надежности производственных линий.

В случаях, когда возможно получение деградационных данных, характеризующих процесс ухудшения работы оборудования, целесообразным становится использование методов анализа надежности на основе деградационных моделей [1]. Для ускорения процесса сбора данных о деградации часто используют комбинированный подход, при

котором данные о деградации собираются в ходе ускоренных испытаний [2].

В литературе представлены различные подходы к построению оптимальных планов экспериментов для определения оптимальных нагрузок и количества исследуемых объектов. Эти методы учитывают ограничения как на допустимые нагрузки, так и на стоимость и продолжительность эксперимента. В предыдущих работах был рассмотрен подход к оптимальному планированию эксперимента, основанный на оптимизации функционала от информационной матрицы Фишера [3], [4]. Алгоритм, предложенный в [5] и [6], позволяет определить оптимальные уровни нагрузок, количество тестируемых объектов и моменты времени для измерения деградационного показателя с учётом критического уровня деградации.

Настоящая работа посвящена построению оптимального плана эксперимента на основе винеровской деградационной модели с ковариатами для анализа надёжности таблетпресса GZP-32.

Постановка задачи

На рис.1 представлен высокоскоростной таблеточный пресс GZP-32, который выполняет весь процесс производства от подачи сырья, наполнения, дозировки, предварительного и основного прессования до вывода готовой продукции. Таблеточный пресс используется в фармакологии, химической и пищевой промышленности.

В июле 2023 года было зафиксировано восемь поломок пуансона, каждая из которых обходится предприятию в сумму около 25000 рублей. В рамках исследования были собраны данные о работе оборудования за месяц для последующего анализа с целью улучшения его надёжности и предотвращения поломок в условиях повышенной влажности и постоянной эксплуатации.



Рис.1 – Таблетпресс GZP-32

Процесс эксплуатации оборудования происходит следующим образом: ежедневно в начале рабочей смены сотрудник загружает материал для производства таблеток и запускает оборудование. В случае необходимости материал дозагружается в течение смены без полной остановки и очистки оборудования, что приводит к накоплению загрязнений и износа. Полная очистка оборудования производится только в конце смены, что возвращает его к первоначальному рабочему состоянию. В летний период, когда влажность воздуха повышается, оборудование требует более частой очистки. При несвоевременном обслуживании повышается риск поломки ключевых элементов, таких как пуансон.

В рамках исследования был установлен мониторинг оборудования с помощью датчиков вибрации и тока, что позволило отслеживать состояние оборудования в реальном времени и выявлять признаки деградации его элементов. А также датчики температуры и влажности, которые позволяют собирать информацию об окружающей среде.

На рис. 2 представлены графики изменения вибрации и силы тока в течение одной смены. Если вибрация близка к нулю, то оборудование остановлено для дозагрузки материала. Чем больше загрязнение оборудования, тем выше сила тока.

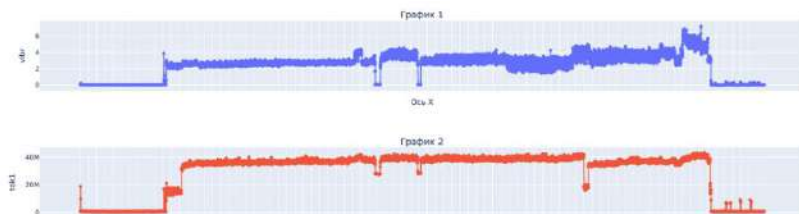


Рис.2 – Графики, отображающие изменения вибрации (синий) и силы тока (красный)

Для анализа деградационных данных, необходимо в первую очередь устранить промежутки простоя оборудования на основании данных о вибрации и произвести объединение соответствующих промежутков измерения силы тока, т.к. дозагрузка материала не влияет на состояние оборудования. После преобразования данных шкала по времени носит неинформативный характер, так как не отражает фактическое время наработки оборудования до отказа. В этом случае предлагается использовать альтернативный вариант – количество килограммов материала, отработанного таблетпрессом до отказа. В среднем в течение часа перерабатывается 50 кг.

Кроме того, для построения винеровской деградационной модели все данные необходимо центрировать путем оценивания процента изменения силы тока относительно начального значения. За критическое значение показателя деградации примем изменение силы тока на 40 %.

Деградационный процесс, полученный после проделанных преобразований, выглядит следующим образом (рис. 3).

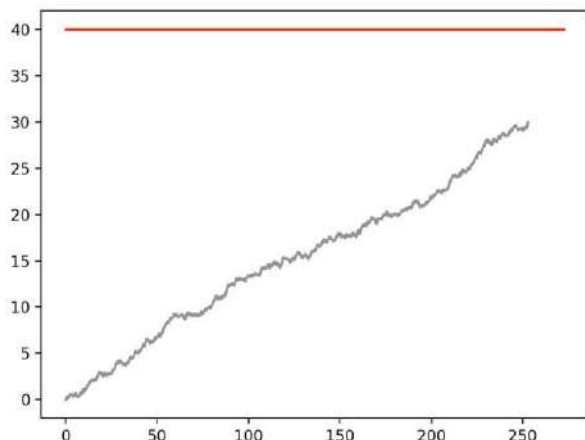


Рис.3 – Преобразованный деградационный процесс

Анализ надежности Таблетпресса GZP-32

Рассмотрим винеровскую деградационную модель с линейной функцией тренда $\rho(t) = t$ при различных функциях от ковариат:

- логарифмическая модель: $r_1(x, \beta) = \exp(\beta \cdot x)$;
- степенная модель: $r_2(x, \beta) = x^\beta = \exp(\beta \cdot \ln(x))$, $x > 0$;
- функция Аррениуса: $r_3(x, \beta) = \exp(\beta / x)$.

Оценки максимального правдоподобия неизвестных параметров для рассмотренных винеровских моделей деградации, а также полученные значения информационных критериев AIC и BIC приведены в таблице 1.

Таблица 1 – Результаты оценивания неизвестных параметров (σ, μ, β) винеровской деградационной модели с линейной функцией тренда

Функция от ковариат	Логарифмическая	Степенная	Аррениуса
Оценки параметров	[0,138; 0,049; -0,033]	[0,286; 0,0008; -1,46]	,987; 1,020; 63.419]
AIC	126	130	128
BIC	208	212	210

Как видно из таблицы 1 винеровская деградационная модель с линейной функцией тренда, где влияние силы тока описывается степенной ковариационной функцией оказалась более предпочтительной моделью по результатам исследований:

$$f(u, \theta_1, \theta_2) = \frac{1}{\sigma \sqrt{2\pi(t/e^{x\beta})}} \exp \left(-\frac{(u - \mu(t/e^{x\beta}))^2}{2\sigma^2(t/e^{x\beta})} \right),$$

с параметрами $\hat{\sigma} = 0,286$, $\hat{\mu} = 0,0008$, $\hat{\beta} = -1,46$.

На рис. 4 представлены графики функций тренда и данные деградации оборудования за 20 рабочих смен. Зеленым цветом выделены 8 деградационных процессов при влажности в помещении 70%, синим цветом при влажности 30%. Красной сплошной линией отмечено критическое значение показателя деградации. Красной пунктирной линией отображены функции тренда для деградационных процессов при соответствующих показателях влажности в помещении.

В результате проведенных исследований было установлено, что при повышенной влажности оборудование требует полной очистки в 4 раза чаще. Однако ресурсы инженеров ограничены, что делает необходимым внедрение более системного подхода к решению этой проблемы. Мы предлагаем поддерживать оптимальный уровень влажности в помещениях, оснащая их специализированными осушителями. Это не только снизит потребность в частой чистке оборудования, но и поможет продлить срок его службы.

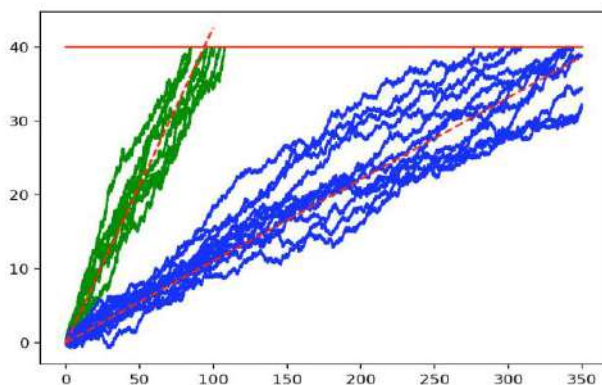


Рис. 4 – Деградационные процессы, основанные на данных о деградации Таблетпресс GZP-32

Используя методы оптимального планирования эксперимента, описанные в [6], подберем оптимальные условия для повторного проведения эксперимента с целью уточнения параметров модели.

В таблице 2 приведены исходный план и полученный D-оптимальный план, а также соответствующие значения определителя ИМФ и УИМФ.

Таблица 2 – Результаты построения D-оптимального плана эксперимента

	Начальный план	Оптимальный план
План	$\xi^0 = \begin{Bmatrix} 30 & 70 \\ 0.6 & 0.4 \end{Bmatrix}$	$\xi = \begin{Bmatrix} 36 & 55 \\ 0.5 & 0.5 \end{Bmatrix}$
ИМФ	3.96e+21	2.75e+21
УИМФ	1.38e+21	1.65e+21

Заключение

В данной статье рассмотрен пример построение анализа надежности для оборудования таблетпресс GZP-32. В ходе исследования было показано, что при повышенной влажности оборудование контроля со стороны оператора каждые 75 кг., в то время, когда более системный подход, а именно оснащение помещения специализированные осушителями позволит сохранить целевое количество наработок таблеток за смену и снизить риск поломки пуансона. Рекомендуется с целью уточнения параметров модели провести повторный эксперимент с настройками, соответствующими оптимальному плану эксперимента.

Литература:

1. Nima G., Lin M., Murthy M., Prasad Y., Yong S. A review on degradation models in reliability analysis. [Proceedings of the 4th World Congress on Engineering Asset Management]. Athens, Greece. – 2009. – P. 379–384. – doi: 10.1007/978-0-85729-320-6_42.
2. Wang C.; Hu Q.; Yu D. Wiener processes with random initial degradation values for step-stress accelerated degradation tests data. [2020 Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling (APARM)]. Vancouver, BC, Canada. – 2020. – doi: 10.1109/APARM49247.2020.9209514.
3. Osintseva E.A., Chimitova E.V. Informacionnaya matrica Fishera dlya vinerovskoj degradacionnoj modeli s uchetom ob'yasnyayushchih peremennyh. [Fisher's information matrix for the Wiener degradation model with covariate] Information Processing and Mathematical Modeling [Proc. of

the Russian scientific and technical conference]. – Novosibirsk. – 2019. – P.92–97. (In Russian).

4. Osintseva, E. Chimitova E. Optimal Design of reliability experiment based on the Wiener degradation model with covariates. Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naja tehnika i informatika. – 2022. – P. 23–33 – doi: 10.17223/19988605/59/3.

5. Osintseva, E. Chimitova E. Optimal design of reliability experiment based on the Wiener degradation model under limitation for the degradation index. 15th International Conference “Intelligent Systems”. – 2023. – Vol: 33(1). – P.64. – URL: <https://doi.org/10.3390/engproc2023033064>.

6. Осинцева, Е.А., Чимитова, Е.В. Оптимальное планирование эксперимента на основе винеровской деградационной модели с ковариатами с использованием условной информационной матрицы Фишера // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2024. – № 67. – С. 9–19. – doi: 10.17223/19988605/67/9.

МОДИФИКАЦИЯ KALMANNET ДЛЯ СЛУЧАЯ НЕОПРЕДЕЛЕННОСТИ В ЗАДАНИИ КОВАРИАЦИОННОЙ МАТРИЦЫ ШУМА ИЗМЕРЕНИЙ

А.В. Комарова

**Новосибирский государственный технический университет,
г. Новосибирск, komnas04@gmail.com**

Научный руководитель: Чубич В.М., д.т.н., профессор

Разработана адаптивная модификация нейросетевого фильтра KalmanNet, которая дополняет классический KalmanNet четвертым рекуррентным блоком GRU, отвечающим за корректировку значений ковариационной матрицы шума измерений. На примере прикладной задачи продемонстрировано преимущество представленной модификации над исходным вариантом KalmanNet и одним из наиболее эффективных современных адаптивных фильтров.

An adaptive modification of the KalmanNet neural network filter has been developed, which complements the classic KalmanNet with a fourth recurrent GRU block responsible for correcting the values of the covariance matrix of measurement noise. Using the example of an applied problem, the advantage of the presented modification over the original version of KalmanNet and one of the most effective modern adaptive filters is demonstrated.

В последнее время наблюдается значительный успех применения глубоких нейронных сетей. Выявлено, что модели, построенные с использованием машинного обучения, могут улавливать важные детали

реального процесса, что позволяет избежать необходимости точно описывать интересующую модель.

В [1] был представлен нейросетевой фильтр KalmanNet, основанный на соответствующем фильтре, но с замененным в нём блоком, отвечающим за вычисление матричного коэффициента усиления Калмана, на рекуррентную нейронную сеть определённой архитектуры. Будучи альтернативой существующим адаптивным фильтрам, KalmanNet зарекомендовал себя как достаточно перспективный подход (см. <https://www.researchgate.net/publication/359180132>). В процессе тестирования предложенной модели на различных задачах было выявлено, что KalmanNet не всегда справляется с ситуациями, когда ковариационная матрица шумов измерений задана неточно. В связи с этим, автором доклада была разработана представленная в статье адаптивная модификация KalmanNet, которая способна справиться с данной проблемой.

Указанная модификация сравнивается с предложенным в [2] адаптивным фильтром AZH (аббревиатура образована по фамилиям её авторов (Akhlaghi, Zhou, Huang)). В настоящий момент данный фильтр является одним из наиболее эффективных и популярных (см. <https://www.researchgate.net/publication/313143778>).

На рис. 1 представлены основные этапы и соответствующие им вычислительные блоки для стандартного фильтра Калмана, AZH и KalmanNet, выделенные разными цветами для облегчения анализа.

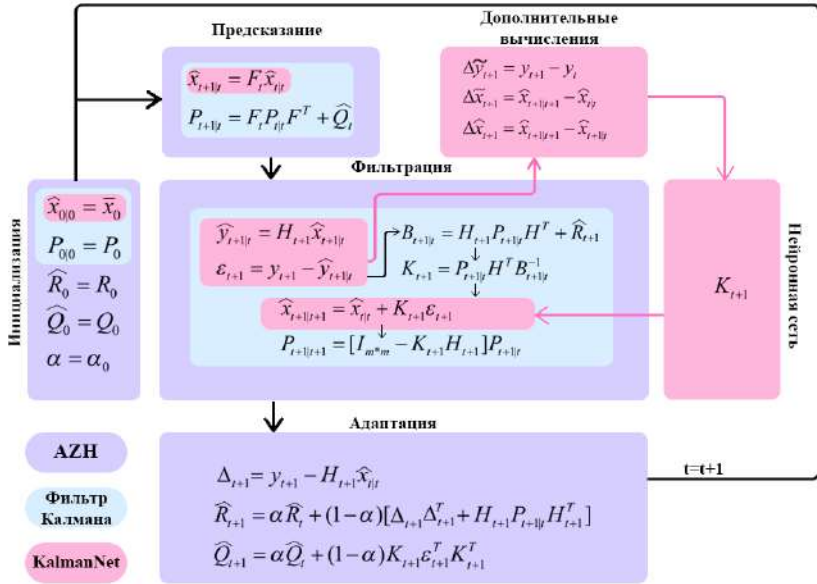


Рис.1 – Структурная схема алгоритмов фильтрации (AZN, Фильтр Калмана, KalmanNet)

Ссылка (https://github.com/KalmanNet/KalmanNet_TSP) на исходный код KalmanNet была представлена в работе [1]. Из него был взят файл KalmanNet_nn.py, содержащий класс с самой структурой KalmanNet, и написано соответствующее обучение.

Суть предлагаемой адаптивной модификации состоит в расширении блока нейронной сети KalmanNet посредством введения дополнительного четвертого GRU, отвечающего за неточности, допущенные в задании ковариационной матрицы шума измерений R_{t+1} .

В результате получена новая архитектура нейронной сети, представленная на рис. 2.

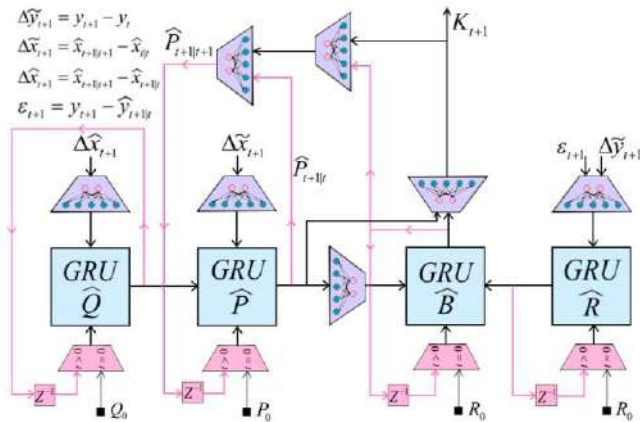


Рис. 2. Структура рекуррентной нейронной сети в адаптивной модификации KalmanNet

Для сравнения разработанной адаптивной модификации KalmanNet и с её исходным вариантом и адаптивным фильтром AZH рассмотрим следующую дискретизованную модель прямолинейного равноускоренного движения материальной точки.

$$x_{t+1} = \begin{pmatrix} 1 & \Delta t & \Delta t^2 / 2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{pmatrix} x_t + w_t; \quad y_{t+1} = (0, 1, 0) x_{t+1} + v_{t+1}, \quad t = 0, 1, \dots, N-1$$

Первая компонента вектора состояния отвечает за перемещение, вторая за скорость и третья за ускорение материальной точки.

Следуя [3], будем считать, что $\Delta t = 0.1$ с., $R = R_0 = R_{\text{исм}} = 0.01$, $Q = Q_0 = Q_{\text{исм}} = \text{diag}(0.01, 0.01, 0.01)$, $\bar{x}_0 = (0, 1, 0)^T$, $P_0 = \text{diag}(0.01, 0.01, 0.01)$. Для данной задачи было найдено оптимальное значение параметра адаптивного фильтра AZH $\alpha_0 = 0.7$.

Оценивать точность предсказаний будем с помощью накопленной среднеквадратичной ошибки ARMSE, вычисляемой

$$\text{ARMSE} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{t=0}^{N-1} (y_t^i - \hat{y}_{t+1|t+1}^i)^2},$$

где M – общее число запусков системы, N – длина каждой выборки. Выберем $M = 100$ и $N = 350$ (было установлено, что при данном значении N процесс обучения проходит наиболее эффективно и быстро).

Проведем сравнительный анализ AZH, KalmanNet и его адаптивной модификации. Результаты представим в таблицах 1 - 3.

Таблица 1 – Значения ARMSE для AZH при различных N, Q и R

N	$Q = Q_0 = Q_{ист}$ $R = R_0 = R_{ист}$	$Q = Q_0 = \frac{Q_{ист}}{100}$ $R = R_0 = \frac{R_{ист}}{100}$	$Q = Q_0 = \frac{Q_{ист}}{100}$ $R = R_0 = R_{ист}$	$Q = Q_0 = Q_{ист}$ $R = R_0 = \frac{R_{ист}}{100}$
250	0.00646	0.6282	0.4628	0.4683
500	0.00625	0.6033	0.4383	0.4299
1000	0.00625	0.6149	0.4366	0.4438

Таблица 2 – Значения ARMSE для KalmanNet при различных N, Q и R

N	$Q = Q_0 = Q_{ист}$ $R = R_0 = R_{ист}$	$Q = Q_0 = \frac{Q_{ист}}{100}$ $R = R_0 = \frac{R_{ист}}{100}$	$Q = Q_0 = \frac{Q_{ист}}{100}$ $R = R_0 = R_{ист}$	$Q = Q_0 = Q_{ист}$ $R = R_0 = \frac{R_{ист}}{100}$
250	0.00473	0.4940	0.3393	0.4479
500	0.00415	0.4400	0.3276	0.4260
1000	0.00433	0.4613	0.3430	0.4464

Таблица 3 – Значения ARMSE для адаптивной модификации KalmanNet при различных N, Q и R

N	$Q = Q_0 = Q_{ист}$ $R = R_0 = R_{ист}$	$Q = Q_0 = \frac{Q_{ист}}{100}$ $R = R_0 = \frac{R_{ист}}{100}$	$Q = Q_0 = Q_{ист}$ $R = R_0 = R_{ист}$	$Q = Q_0 = Q_{ист}$ $R = R_0 = \frac{R_{ист}}{100}$
250	0.00359	0.4333	0.2926	0.2620
500	0.00347	0.4086	0.2634	0.3041
1000	0.00353	0.3865	0.2776	0.3733

Из таблиц 1 – 3 следует, что при всех рассматриваемых парах начальных значений Q_0 и R_0 и выбранных объемах тестовых выборок разработанная адаптивная модификация KalmanNet имеет наименьшие значения ARMSE. Особенно преимущество модификации KalmanNet заметно для случая, когда $Q = Q_0 = Q_{ист}$ и $R = R_0 = R_{ист} / 100$. Здесь KalmanNet показал сопоставимые результаты с AZH, в отличие от разработанной адаптивной модификации, для которой ARMSE заметно ниже.

Будущие исследования могут быть направлены на расширение области применения разработанной адаптивной модификации

KalmanNet на более сложные, в том числе нелинейные, модели, а также на решение задач практического характера.

Литература:

1. Revach N.G., Shlezinger R.J.G., van Sloun Y.C. Eldar KalmanNet: Data-driven Kalman filtering // IEEE Int. Conf. on Acoustics, Speech and Signal Processing. – Toronto, 2021. – P. 3905–3909.
2. Akhlaghi S., Zhou N., Huang Z. Adaptive Adjustment of Noise Covariance in Kalman Filter for Dynamic State Estimation // IEEE Power and Energy Conf. Gen. Meeting. – Chicago, 2017. – P. 1-5.
3. Прокофьева А.Э., Чубич В.М. Сравнительный анализ некоторых робастных фильтров для нестационарных линейных дискретных систем // Вестник ИргТУ. – 2017. – Т. 21, № 12. – С. 123-137.

ИССЛЕДОВАНИЕ ДИСКРЕТНОЙ МОДЕЛИ «PARTICAL LIFE»

Д.А. Пляскин

Томский государственный университет систем управления
и радиоэлектроники,

г. Томск, zelcamn2@yandex.ru

Научный руководитель: Костелей Я.В., к.т.н, доцент

В статье рассматривается модель «Жизнь Частиц» схожая с игрой «Жизнь» Джона Конвея. В отличие от неё «Жизнь Частиц» оперирует не сеткой и «соседями», а частицами и силами, которые возникают между ними. Несмотря на это, данная система всё ещё, имея простые правила, реализует нетривиальное поведение.

This paper reviews a model “Particle Life” that is like John Conway’s “Game of Life”. In contrast to it “Particle Life” does not use grid and “neighbors”, but particles and forces between them. However, with a simple rules it still has nontrivial behavior.

Введение. В ходе развития информационных технологий появляются новые методы моделирования различных систем. Особенный интерес вызывают системы, обладающие нетривиальным поведением, которое трудно предсказать, не запустив программу.

Такие системы, формируемые из децентрализованной самоорганизации множества локальных взаимодействий, называются искусственной жизнью [1]. Искусственная жизнь используется учёными для изучения живых систем и их эволюции.

Однако эти системы вызывают интерес не только среди исследователей. В настоящее время модификации некоторых систем

становятся популярными среди программистов и цифровых художников. Первых заинтересовывает простота реализации систем на различных языках программирования. Вторых – эстетическая сторона создаваемых симуляциями образов.

Основной целью данной работы является изучение модели «Particle Life» и анализ её возможного использования в творческой деятельности.

Для достижения указанных целей необходимо:

1. Проанализировать существующие реализации указанной модели.
2. Разработать модель в виде компьютерного приложения.
3. Рассмотреть различные области применения модели.

Реализации алгоритмов. В 2015 г. Джеффри Вентрелла опубликовал видео с описанием модели симуляции взаимодействующих частиц «Clusters» [2]. Правила, описанные в ней, послужили основой для рассматриваемой модели.

Главным образом они включают в себя описание объектов «частиц», имеющих следующие параметры: координаты в пространстве, скорость, ускорение, цвет. Эти объекты взаимодействуют друг с другом, исходя из матриц отношений, которые делят их по цвету. Матрицы определяют, будут ли частицы притягиваться или отталкиваться друг от друга, на каком расстоянии это будет происходить и с какой силой.

В 2022 г. пользователь интернета с никнеймом «Tom Mohr» выложил видео, объясняющее данную модель, которую он, модифицировав, назвал «Particle Life». Правила его версии ничем не отличались от версии Джеффри Вентрелла, однако усовершенствованный программный код позволилкратно увеличить число симулируемых частиц.

Разработка. Современное программное обеспечение подразумевает наличие архитектуры, разделяющей различные компоненты приложения. Разработанное приложение разделено на код графической составляющей, который выводит результаты симуляции на экран, и код, непосредственно реализующий модель. В качестве инструментов выбраны язык C++, библиотека создания окон GLFW, библиотека создания пользовательского интерфейса ImGUI и API (программный интерфейс) OpenGL.

Графическая часть. Джеффри Вентрелла и «Tom Mohr» используют одинаковый подход для визуализации данных. Частицы представлены окружностями с одинаковым радиусом и цветом. В модели «Particle Life» пространство симуляции не замкнуто по всем его сторонам. Если частица выйдет за нижнюю часть экрана, то она

появится в его верхней части. То же самое верно для левой и правой границ.

Для создания окна и обращения к графическому API использовался GLFW. Во время генерации окружности в вертексный шейдер передаются координата центра окружности и цвет частицы. Далее с помощью геометрического шейдера генерируются дополнительные вершины для создания «окружности» из треугольников (рисунок 1).

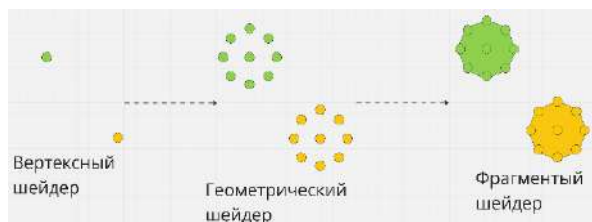


Рис. 1 – Процесс отрисовки окружности

Симуляционная часть. Описанные ранее правила моделирования применялись при разработке симуляционной части приложения. Нововведением является дополнительный способ расчёта воздействия сил, представленный в видеоблоге «Programming Chaos» [3].

Если коэффициент в матрице отношений между двумя частицами x и y положительный для первого, то базовая функция расчёта силы при достижении расстояния $dist$ начнёт притягивать частицу x к частице y . Сила будет линейно возрастать до расстояния $radius$, находящегося посередине расстояния воздействия силы, и убывать при отдалении от него. Когда частица x пересечёт расстояние $minDistance$, сила станет отрицательной и начнёт отталкивать её от частицы y . Частица x будет постоянно находиться на некотором удалении от частицы y , то притягиваясь к ней, то отталкиваясь от неё.

Автор канала «Programming Chaos», утверждая, что такая функция «немного» сложна в реализации и не соответствует реальным физическим силам, предложил изменённую версию. В ней компонент «притягивания» рассчитывается всегда, линейно убывая до расстояния $radius$. Компонент «отталкивания» добавляется только при достижении частицы x расстояния $minDistance$. На рисунке 2 показаны описанные функции. Отдельной обработки случая, когда частица отталкивается, а не притягивается к другой, не требуется, так как получившиеся результаты умножаются на коэффициент в матрице.

Результат реализации системы представлен на рисунке 3.

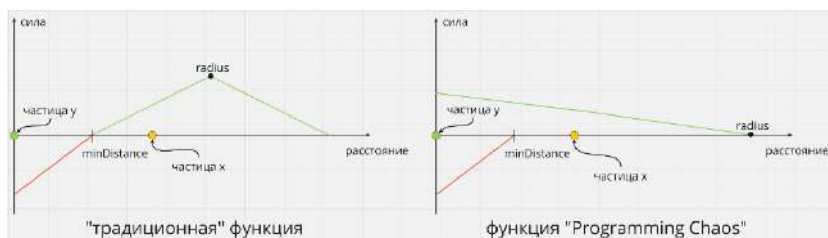


Рис. 2 – Функции силы, используемые для моделирования

Области применения. Ранее было сказано, что, данная модель легка в реализации, соответственно, её можно использовать при обучении программированию и методам оптимизации программного обеспечения. Второе является валидным, так как модель в своём базовом исполнении неспособна осуществлять симуляцию большого количества частиц. Для этого программисту нужно найти пути увеличения производительности. Например, использованием kd-деревьев для поиска соседних частиц.

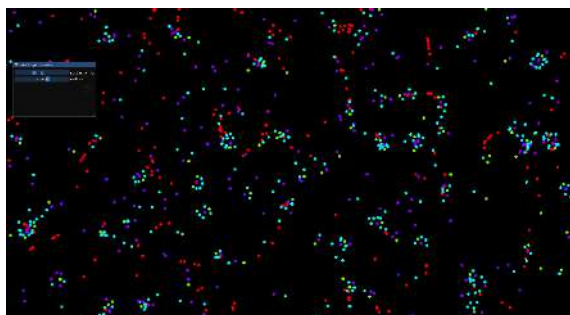


Рис. 3 – Результат разработки

Рассматриваемая модель также набирает популярность среди цифровых художников. Они, экспериментируя с комбинациями матриц отношений, выявляют неожиданные композиции. Некоторые создают с помощью «Particle Life» не только картины и анимации, но и шейдеры для видеоигр. Примеры представлены на рисунке 4.



Рис. 4 – Примеры использования «Particle Life» в искусстве

Результаты. В ходе данной работы был проведён анализ модели «Particle Life», который включал рассмотрение существующих реализаций, её разработку и исследование возможного использования. Выяснено, что данная модель была разработана относительно недавно (2015 г.), её реализация проста, а области использования затрагивают не только задачи обучения программированию, но и цифровое искусство.

Литература:

1. Галкин Д.В. Искусственная жизнь: наука и компьютерные технологии в современном искусстве // Вестник Томского государственного университета. – 2011. – №350. – С. 74–79.
2. Jeffrey Ventrella. Explains Clusters [Электронный ресурс] // Vimeo.com – URL: <https://vimeo.com/222974687> (дата обращения 09.11.2024).
3. Programming Particle Life [Электронный ресурс] // Youtube.com – URL: https://youtu.be/xiUpAeos168?si=_1LK0_RylJfgkoeO (дата обращения 09.11.2024).

ОЦЕНКА ЭФФЕКТИВНОСТИ СПОСОБОВ ХРАНЕНИЯ СТРУКТУР ДЕРЕВЬЕВ И/ИЛИ ИХ ВАРИАНТОВ

К.С. Ражев, А.В. Токарева

**Томский государственный университет систем управления
и радиоэлектроники,**

г. Томск, alina.v.tokareva@tusur.ru

Научный руководитель: Токарева А.В., м.н.с. ЛАТИДС

Данная работа посвящена проблеме хранения структур деревьев И/ИЛИ и их вариантов. Рассмотрены существующие способы, позволяющие хранить

соответствующие структуры данных в памяти компьютера, и выполнена их программная реализация. Также проведены вычислительные эксперименты для оценки эффективности с точки зрения скорости обработки данных.

This paper is devoted to the problem of storing the structures of AND/OR trees and their variants. The existing methods allowing storing the corresponding data structures in the computer memory are considered, and their software implementation is performed. Computational experiments are also conducted to evaluate the efficiency in terms of data processing speed.

С динамическим ростом объема используемой информации актуальной становится задача выбора наиболее подходящего способа хранения данных. Одной из широко применяемых структур данных является древовидная структура. Деревья популярны для решения задач, связанных с организацией иерархических взаимосвязей, разработкой классификаторов и каталогов [1-3]. Разнообразие типов древовидных структур подтверждает их важность для решения частных задач. Данная работа посвящена исследованию структур деревьев И/ИЛИ. Деревья И/ИЛИ – это древовидная структура представления множества с узлами двух типов: И-узел (подразумевающий операцию умножения дочерних узлов) и ИЛИ-узел (подразумевающий операцию сложения дочерних узлов).

С помощью структур деревьев И/ИЛИ представляется возможным решение таких задач как построение многовариантных тестовых заданий [4], кодирование текстовой информации [5], создание систем идентификации и прослеживаемости изделий [6]. Однако при разработке соответствующего программного обеспечения возникает потребность выбора оптимального способа хранения таких структур. Целью данной работы является экспериментальное исследование эффективности способов хранения структур деревьев И/ИЛИ и их вариантов в памяти компьютера.

В рамках эксперимента были рассмотрены следующие способы хранения деревьев И/ИЛИ:

- хранение на основе массивов, где каждому элементу массива сопоставляется конкретный узел дерева и хранится указатель на родительский узел;
- сериализация объектов с помощью JSON (JavaScript Object Notation), позволяющая хранить массив пар «ключ:значение»;
- правильная скобочная последовательность, в которой каждому узлу дерева сопоставляется пара открывающей и закрывающей скобок, а внутреннее содержание скобок соответствует описанию узлов-потомков;

- подход к реализации биективного отображения структуры дерева И/ИЛИ и его вариантов на схему реляционной базы данных, предложенный в работе [7], в основе которого лежит представление дерева И/ИЛИ в виде взаимосвязанных таблиц в реляционной базе данных с выделением родителей и потомков.

Для экспериментов было построено дерево И/ИЛИ, схематичное изображение представлено на рисунке 1.

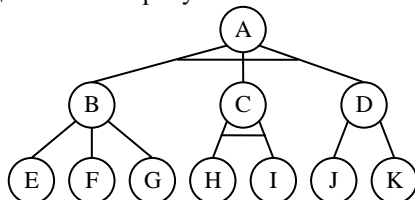


Рис. 1 – Схематичное изображение дерева И/ИЛИ для эксперимента

Для проведения экспериментов по скорости работы алгоритмов обработки была написана программа на языке C++. Дерево И/ИЛИ было представлено в программе с помощью скобочной структуры, массивов и сериализации с помощью JSON. Было выполнено измерение времени загрузки данных из постоянной памяти в оперативную память, для минимизации случайных погрешностей проведено усреднение результатов 500 повторений одного испытания. Испытание представляло собой последовательное открытие файла, парсинг строки и запись дерева И/ИЛИ в память программы. Далее были произведены измерения среднего времени преобразования и сохранения варианта дерева из ОЗУ в ПЗУ. Среднее время загрузки данных в оперативную память из постоянной представлено в таблице 1.

Таблица 1 – Время загрузки дерева И/ИЛИ из постоянной в оперативную память

Способ хранения	Среднее время загрузки в ОЗУ, наносекунд	Среднее время сохранения в ПЗУ, наносекунд
Скобочное представление	67 303	234 168
На основе массивов	91 956	237 764
Сериализация в JSON	218 196	262 515

Полученные результаты экспериментов позволяют сделать вывод о том, что скобочное представление является наиболее лучшим с точки зрения скорости обработки небольшого по количеству содержимой информации дерева И/ИЛИ. Однако предполагается, что для хранения

больших деревьев И/ИЛИ и их вариантов более подходящим вариантом является подход с реализацией биективного отображения дерева И/ИЛИ на схему реляционной базы данных. Для проведения дальнейшего эксперимента дерево И/ИЛИ было представлено в базе данных согласно описанному подходу [7] (рисунок 3).

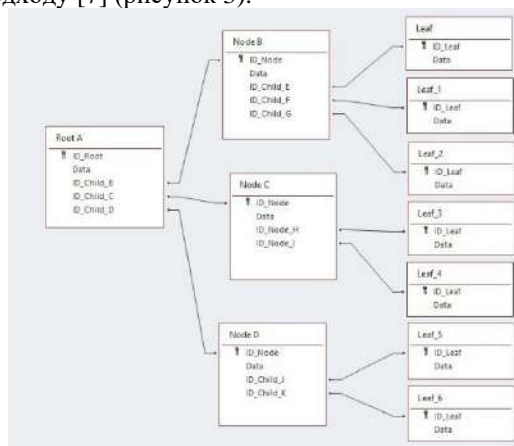


Рис. 3 – Биективное отображение дерева И/ИЛИ в базу данных

В качестве эксперимента реализовывалась задача поиска метки по структуре данных. В случаях с приведенными выше способами происходило последовательное открытие файла в постоянной памяти и замерялось время парсинга всего содержимого со всеми записанными вариантами в файле до момента нахождения искомой метки. В базу данных отправлялись SQL-запросы на сравнение полей «Data» в каждой таблице со значением метки для поиска, с остановкой поиска после успешного обнаружения нужной таблицы. Было проведено две серии экспериментов, в первой серии искомая метка находилась в первой записи, во второй серии – в последней записи.

Эксперименты проводились поочередно с увеличением количества символов в метках узлов. Полученные результаты записаны в таблице 2.

**Таблица 2 – Время поиска по вариантам деревьев И/ИЛИ
различного положения и размера**

Способ хранения	Положение записи с искомой меткой	Размер дерева (примерное суммарное количество символов) и время поиска, наносекунд		
		40	250	1000
Реляционная база данных	Первая	263 372	269 624	288 458
	Последняя	1 316 860	1 348 120	1 442 290
Скобочное представление	Первая	64 900	79 313	130 793
	Последняя	389 400	475 878	784 758
На основе массивов	Первая	136 759	246 102	636 612
	Последняя	820 554	1 476 612	3 819 672

Таким образом, проведённые эксперименты показали, что при обработке деревьев И/ИЛИ больших размеров наиболее эффективным по времени обработки данных является подход биективного отображения дерева И/ИЛИ на схему реляционной базы данных. В дальнейшем данный подход будет использован при разработке информационных систем, в которых используются деревья И/ИЛИ.

Работа выполнена при финансовой поддержке РФН в рамках проекта № 22-71-10052.

Литература:

1. *Subero A.* Codeless data structures and algorithms. – USA: Springer, 2020. – 143 p.
2. *Celko J.* Trees and hierarchies in SQL for smarties. – USA: Morgan Kaufmann Publishers, 2012. – 296 p.
3. *Шаров В.Ю., Горшкова Ю.Ю., Филоненко И.Н.* Сохранение древовидных структур в базе данных и управление ими // Комплексные проблемы развития науки, образования и экономики региона. – 2015. – № 2. – С. 223–231.
4. *Зорин Ю.А., Посов И.А.* Инструментальные системы построения и получения многовариантных тестовых заданий // Компьютерные инструменты в образовании. – 2014. – № 1. – С. 14–25.
5. *Кручинин В.В., Люшкин Б.А.* Метод кодирования информационных объектов на основе деревьев И/ИЛИ // Доклады ТУСУР. — 2010. – № 1. – С. 170–172.
6. *Токарева А.В., Кручинин Д.В.* Модификация метода идентификации и прослеживаемости сложных технических изделий с применением алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ // Вестник СибГУТИ. – 2024. – № 3. – С. 73–85.

7. *Шабля Ю.В., Токарева А.В.* Способы хранения структур деревьев И/ИЛИ и их вариантов в оперативной и постоянной компьютерной памяти // Доклады ТУСУР. – 2024. – № 2. – С. 44–50.

ОПРЕДЕЛЕНИЕ КОЭФФИЦИЕНТА ОСЛАБЛЕНИЯ ПРИ ОБРАБОТКЕ ДАННЫХ ЛИДАРНОГО ЗОНДИРОВАНИЯ

В.Р. Потапова¹, Н.В. Дьяченко², Е.Ю. Михтеева²

**¹Балтийский государственный технический университет
«ВОЕНМЕХ» им. Д.Ф. Устинова,
г. Санкт-Петербург, veronikar05@yandex.ru**

**²Российский государственный гидрометеорологический
университет, г. Санкт-Петербург.**

В данной работе рассматривается задача определения коэффициента ослабления, заключающаяся в нахождении неизвестной постоянной σ с минимальной погрешностью. Для достижения этой цели обычно минимизируют сумму определенных величин, что приводит к получению равенства для расчета коэффициента ослабления. Однако существующий метод имеет недостаток, связанный с погрешностью численного интегрирования. В предложенном исследовании представлен новый метод, который устраняет данную проблему. Также рассматривается система уравнений, позволяющая более точно определить коэффициент ослабления.

In this paper, the problem of determining the attenuation coefficient is considered, which consists in finding an unknown constant σ with a minimum error. To achieve this goal, it is necessary to minimize the sum of certain values, which leads to equality for calculating the attenuation coefficient. However, the existing method has a disadvantage associated with the error of numerical integration. The proposed study presents a new method that eliminates this problem. A system of equations is also considered, which makes it possible to more accurately determine the attenuation coefficient.

Для анализа атмосферных загрязнений можно применять технологию ЛИДАР (от англ. Light Detection and Ranging), которая позволяет измерять концентрацию аэрозольных частиц и газов. ЛИДАР функционирует на основе активных оптических систем, использующих явления отражения и рассеивания света в прозрачных и полупрозрачных средах для получения информации об удаленных объектах. Для обработки данных, полученных с помощью ЛИДАРа, применяется эхолокационное уравнение [1 - 7] которое устанавливает связь между сигналом обратного рассеяния и коэффициентами ослабления и обратного рассеяния.

$$P_r(R) = P_0 * \left(\frac{c\tau}{2}\right) * \beta(R) * A_r * R - \exp(-2 \int \alpha(r) dr) \quad (1)$$

где

$P_r(R)$ – мгновенное значение принимаемой мощности в момент времени t ,

P_0 – посылаемая мощность в момент времени t_0 ,

c – скорость света,

τ – длительность импульса,

β – объемный коэффициент рассеяния,

R – расстояние до объекта,

A_r – эффективная площадь приемника,

α – объемный коэффициент ослабления.

Для обработки данных, полученных с помощью ЛИДАРа, обычно применяется метод последовательного приближения, который требует значительных временных затрат. Чтобы повысить эффективность этого процесса, можно использовать метод, основанный на итерационных процессах [8 – 11]. Однако из-за нелинейности задачи результаты могут оказаться неточными. Для решения этой проблемы можно применить метод, не требующий итерационного процесса [12]. Если атмосфера достаточно однородна, то есть коэффициенты ослабления и обратного рассеяния остаются постоянными на протяжении всего зондирования, можно записать лидарное уравнение в следующем виде:

$$P_i = P_* + \frac{B}{R_i^2} \exp(-2\sigma R_i), \quad (2)$$

где

$$B = A\beta. \quad (3)$$

A – постоянная лидара,

β – коэффициент обратного рассеяния,

σ – коэффициент ослабления,

P – мощность сигнала обратного рассеяния,

P_* – мощность солнечного излучения, рассеянного атмосферой в направлении на приемное устройство ЛИДАРа,

R – расстояние между лидаром и i – M рассеивающим элементом.

Для того чтобы определить концентрацию загрязняющих веществ, необходимо свести к минимуму погрешность коэффициента ослабления σ . Для достижения этой цели нужно минимизировать сумму:

$$\delta^2 = \sum_{i=1}^n (C - S_i - 2\sigma \int_{R_0}^{R_i} S dr)^2, \quad (4)$$

где $S = (P - P_*)R^2$.

Для вычисления σ , коэффициента ослабления, образуется равенство

$$\sigma = \frac{1}{2} \frac{\sum_{i=1}^n S_i \sum_{j=1}^n \int_{R_0}^{R_i} S dR - n \sum_{i=1}^n (S_i \int_{R_0}^{R_i} S dR)}{n \sum_{i=1}^n (\int_{R_0}^{R_i} S dR)^2 - (\sum_{i=1}^n \int_{R_0}^{R_i} S dR)^2}. \quad (5)$$

В уравнениях (4, 5) присутствует погрешность численного интегрирования, что является недостатком данного метода [13 - 19]. Исследование показало, что обработка данных ЛИДАРa может быть сложной и занимать много времени при использовании метода последовательного приближения [20 - 22]. Для ускорения этого процесса можно применять метод итерационных процессов, однако это может привести к неточным результатам из-за нелинейности задачи. Для решения данной проблемы можно использовать метод, не требующий итерационного процесса.

При использовании нового алгоритма вычисления коэффициента ослабления предлагается рассматривать следующую систему уравнений

$$P_1 = P_* + \frac{B}{R_1^2} \exp(-2\sigma R_1), \quad (6)$$

$$P_2 = P_* + \frac{B}{R_2^2} \exp(-2\sigma R_2), \quad (7)$$

$$P_n = P_* + \frac{B}{R_n^2} \exp(-2\sigma R_n), \quad (8)$$

$$P_{n+1} = P_* + \frac{B}{R_{n+1}^2} \exp(-2\sigma R_{n+1}). \quad (9)$$

При суммировании уравнений системы получаем

$$\sum_{i=1}^n (P_i - P_*) R_i^2 = B \exp(-2\sigma R_1) \{1 + \exp(-2\sigma h) + \dots + \exp(-2\sigma(n-1)h)\} \quad (10)$$

Если обозначить

$$S_n = 1 + \exp(-2\sigma h) + \dots + \exp(-2\sigma(n-1)h) \quad (11)$$

то

$$S_n = \frac{1 - \exp(-2\sigma n h)}{1 - \exp(-2\sigma h)}, \quad (12)$$

причем

$$\sum_1^n (P_i - P_*) R_i^2 = \frac{(P_1 - P_*) R_1^2 - (P_{n+1} - P_*) R_{n+1}^2}{1 - \exp(-2\sigma h)} . \quad (13)$$

Т. о., следует находить минимум суммы

$$\sum_{j=1}^n \left\{ C - (P_{j+1} - P_*) R_{j+1}^2 - \Delta \sum_1^n (P_i - P_*) R_i^2 \right\}, \quad (14)$$

где C – неизвестная константа,

$$\Delta = 1 - \exp(-2\sigma h), \quad (15)$$

поэтому

$$\sigma = -\ln(1 - \Delta)/2h . \quad (16)$$

Для постоянной Δ получаем равенство:

$$\Delta = \frac{\sum_{j=1}^n S_{j+1} \sum_{j=1}^n \sum_{i=1}^j S_i - n \sum_{j=1}^n S_{j+1} \sum_{i=1}^j S_i}{n \sum_{j=1}^n (\sum_{i=1}^j S_i)^2 - (\sum_{j=1}^n \sum_{i=1}^j S_i)^2}, \quad (17)$$

где

$$S_{j+1} = (P_{j+1} - P_*) R_{j+1}^2 \quad (18)$$

$$S_i = (P_i - P_*) R_i^2 \quad (19)$$

Результаты определения коэффициента ослабления для $R_0 = 2,5 \text{ км}$ представлены в таблице 1.

Таблица 1 – Результаты определения фоновой засветки для $R_0 = 2,5 \text{ км}$

$\Delta R, \text{ км}$	2	4	6	8	10
$\sigma \text{ км}^{-1} (5)$	0,066	0,083	0,083	0,098	0,107
$\sigma \text{ км}^{-1} (16), (17)$	0,067	0,084	0,085	0,101	0,115

Анализ представленных данных свидетельствует о том, что приближенный метод обладает достаточно высокой точностью.

В заключение можно отметить, что несмотря на возникшие трудности, методы обработки данных ЛИДАРа все же обеспечивают достаточно точные и быстрые результаты. Тем не менее, необходимо дальнейшее совершенствование и разработка новых методов для обеспечения быстрой обработки данных не только для аэрозольной фракции, но и для концентрации газовых компонентов при лидарном зондировании.

Литература:

1. *Егоров А.Д., Потапова И.А., Ржонсницкая Ю.Б., Драбенко В.А.* Лидарное зондирование атмосферного аэрозоля // Метеорологический вестник. – 2013. – Т. 5, № 2. – С. 42-51.
2. *Егоров А.Д., Потапова И.А.* Лидарные исследования прозрачности атмосферы // Труды Главной геофизической обсерватории им. А.И. Воейкова. – 2004. – № 553. – С. 131-142.
3. *Егоров А.Д., Потапова И.А., Ржонсницкая Ю.Б., Ошуркова А.А.* Методы лидарного зондирования аэрозольных загрязнений атмосферы импульсами конечной длительности // Ученые записки Российского государственного гидрометеорологического университета. – 2011. – № 18. – С. 48-56.
4. *Бобровский А.П., Дьяченко Н.В., Косцов В.В., Михтеева Е.Ю., Потапова И.А., Скобликова А.Л., Хлябич П.П., Яковлева Т.Ю.* Лидарная прозрачнотрия слабо замутненной атмосферы // В сборнике: Современные тенденции и перспективы развития гидрометеорологии в России. Материалы IV Всероссийской научно-практической конференции, приуроченной к Году науки и технологий. Иркутск, 2021. С. 495-500.
5. *Драбенко В.А., Егоров А.Д., Галкин И.А., Потапова И.А., Драбенко Д.В., Игнатенко В.М., Потапова В.Р.* Способ аспирационной оптической спектрометрии аэрозольных частиц // Патент на изобретение RU 2681256 C2, 05.03.2019. Заявка № 2016120373 от 25.05.2016.
6. *Потапова И.А., Егоров А.Д., Дьяченко Н.В., Саноцкая Н.А., Скобликова А.Л., Яковлева Т.Ю.* Интерпретация слабого радиационного излучения при лидарном зондировании атмосферы // В книге: Атмосферная радиация и динамика (МСАРД – 2019). Международный симпозиум. Тезисы, 2019. – С. 189-190.
7. *Дьяченко Н.В., Потапова И.А., Саватннеев А.Н., Хлябич П.П., Яковлева Т.Ю.* Интерпретация лидарной информации при зондировании грубодисперсной аэрозольной фракции // В сборнике: Инновационные методы математики и физики в экологических и

гидрометеорологических исследованиях. Сборник трудов Всероссийской научно-практической конференции. Под редакцией И.В. Зайцевой. Санкт-Петербург, 2023. – С. 105-110.

8. *Егоров А.Д., Потапова И.А., Ржонсницкая Ю.Б.* Обращение слабых сигналов обратного рассеяния // Ученые записки Российского государственного гидрометеорологического университета. – 2011. – № 22. – С. 144-148.

9. *Егоров А.Д., Потапова И.А., Ржонсницкая Ю.Б., Саноцкая Н.А.* Способ дистанционного оптического зондирования слабо рассеивающей атмосферы // Патент на изобретение RU 2495452 С2, 10.10.2013. Заявка № 2011143951/28 от 31.10.2011.

10. *Егоров А.Д., Потапова И.А.* Способ определения прозрачности атмосферы // Патент на изобретение RU 2439626 С2, 10.01.2012. Заявка № 2009144060/28 от 27.11.2009.

11. *Егоров А.Д., Блаkitная П.А., Потапова И.А., Ржонсницкая Ю.Б., Саноцкая Н.А.* Способ оптического зондирования атмосферы // Патент на изобретение RU 2441261 С1, 27.01.2012. Заявка № 2010143286/28 от 22.10.2010.

12. *Бобровский А.П., Дьяченко Н.В., Егоров А.Д., Косцов В.В., Калинин Д.В., Ку克林 О.А., Михтеева Е.Ю., Потапова И.А., Скобликова А.Л., Хлябич П.П., Яковлева Т.Ю.* Алгоритм лидарного определения фоновой засветки и коэффициента ослабления слабо замутненной атмосферы без итерационного процесса // В книге: Современные проблемы гидрометеорологии и устойчивого развития Российской Федерации. Сборник тезисов Всероссийской научно-практической конференции, 2019. – С. 58-60.

13. *Потапова И.А.* Оценка инструментальных погрешностей лидарных измерений атмосферных характеристик // Ученые записки Российского государственного гидрометеорологического университета. – 2009. – № 9. – С. 86-91.

14. *Егоров А.Д., Потапова И.А., Ржонсницкая Ю.Б., Саноцкая Н.А.* Оценка погрешностей результатов лидарного зондирования слабо замутненной атмосферы/ Оптический журнал. 2013. Т. 80. № 3. С. 58-60.

15. *Егоров А.Д., Потапова И.А., Ржонсницкая Ю.Б.* Оценка случайных погрешностей лидарных измерений атмосферных характеристик // Ученые записки Российского государственного гидрометеорологического университета. – 2011. – № 17. – С. 51-55.

16. *Егоров А.Д., Потапова И.А., Привалов Д.В., Ржонсеницкая Ю.Б.* Систематические и случайные погрешности обращения лидарных сигналов малой мощности // Депонированная рукопись № 1230-гм2005 16.12.2005

17. Егоров А.Д., Потапова И.А., Привалов Д.В. Систематические погрешности обращения лидарных сигналов малой мощности // Труды Главной геофизической обсерватории им. А.И. Воейкова. – 2006. – № 555. – С. 30-34.

18. Yegorov A.D., Potapova I.A., Rzhonsnitskaya Yu.B., Sanotskaya N.A., Shchadin A.V. Atmospheric aerosol measurements and reliability problem: new results // International Journal of Remote Sensing, 2014. – Vol. 35. – P. 5750-5765.

19. Potapova I.A., Bobrovsky A.P., Dyachenko N.V., Rzhonsnitskaya Yu.B., Sanotskaya N.A., Mikhiteeva E.Yu., Khlyabich P.P., Kostsov V.V. Method for solving the inverse problem of lidar sounding of a weakly turbid atmosphere // Journal of Instrument Engineering. – 2021. – Т. 64, № 5. – С. 384-391.

20. Потапова И.А., Волкова Н.А., Михтеева Е.Ю., Потапова В.Р., Хлябич П.П. Анализ воздействия естественного излучения атмосферы на принимаемый сигнал при исследовании прозрачности // В сборнике: Инновационные методы математики и физики в экологических и гидрометеорологических исследованиях. сборник трудов Международной научно-практической конференции. Санкт-Петербург, 2024. – С. 160-167.

21. Бобровский А.П., Дьяченко Н.В., Михтеева Е.Ю., Потапова И.А., Потапова В.Р., Скобликова А.Л., Хлябич П.П., Яковлева Т.Ю. Влияние воздействия естественного рассеянного излучения атмосферы на отраженный сигнал при исследовании атмосферы низкой мутности // В сборнике: Гидрометеорология и физика атмосферы: современные достижения и тенденции развития. Материалы II Международной научно-практической конференции. Санкт-Петербург, 2024. – С. 34-39.

22. Махлин Н.К., Потапова И.А., Ягнюков С.А. Обратная задача лидарного зондирования слабо замутненной атмосферы // В сборнике: Инновационные методы математики и физики в экологических и гидрометеорологических исследованиях. Сборник трудов Всероссийской научно-практической конференции. Под редакцией И.В. Зайцевой. Санкт-Петербург, 2023. – С. 129-137.

КВАДРАТНО-КОРНЕВАЯ РОБАСТНАЯ МОДИФИКАЦИЯ АДАПТИВНОГО ФИЛЬТРА АХЛАГИ – ЧЖОУ – ХУАНГА

С.А. Евланова

Новосибирский государственный технический университет,
г. Новосибирск, evlanovasof@yandex.ru

Научный руководитель: Чубич В. М., д.т.н., профессор

Разработана и программно реализована квадратно-корневая робастная модификация адаптивного фильтра Ахлагхи (Akhlaghi) – Чжоу (Zhou) – Хуанга (Huang). На примере модели равноускоренного движения материальной точки проведен сравнительный анализ эффективности работы указанного фильтра с фильтрами Ахлагхи (Akhlaghi) – Чжоу (Zhou) – Хуанга (Huang) и Изанлу (Izanloo) – Фейкуруиана (Fakoorian) – Джазди (Yazdi) – Саймона (Simon).

A square-root robust modification of the Akhlaghi-Zhou-Huang adaptive filter has been developed and implemented in software. Using the example of a model of uniformly accelerated motion of a material point, a comparative analysis of the efficiency of the specified filter with the Akhlaghi-Zhou-Huang and Izanloo-Fakoorian-Yazdi-Simon filters has been carried out.

Рассмотрим следующую модель управляемой и наблюдаемой дискретной линейной системы в пространстве состояний:

$$x(t_{k+1}) = F(t_k)x(t_k) + \Psi(t_k)u(t_k) + \Gamma(t_k)w(t_k),$$

$$y(t_{k+1}) = H(t_{k+1})x(t_{k+1}) + v(t_{k+1}), \quad k = \overline{0, N-1}.$$

Здесь $x(\cdot)$ – вектор состояния; $u(\cdot)$ – вектор управления; $w(\cdot)$ – вектор шума системы; $y(\cdot)$ – вектор измерения; $v(\cdot)$ – вектор шума измерения; $F(\cdot)$ – матрица состояния; $\Psi(\cdot)$ – матрица управления; $\Gamma(\cdot)$ – матрица шумов системы; $H(\cdot)$ – матрица измерения.

Будем считать, что случайные процессы $\{w(t_k), k = \overline{0, N-1}\}$ и $\{v(t_{k+1}), k = \overline{0, N-1}\}$ являются белыми гауссовскими шумами, причем

$$\begin{aligned} E[w(t_k)] &= 0, \quad E[w(t_k)w^T(t_i)] = Q(t_k)\delta_{ki}; & E[v(t_k)] &= 0, \\ E[v(t_{k+1})v^T(t_{i+1})] &= R(t_{k+1})\delta_{ki}; & E[w(t_k)v^T(t_i)] &= 0; \end{aligned}$$

начальное состояние $x(t_0)$ имеет нормальное распределение с математическим ожиданием $\bar{x}(t_0)$, ковариационной матрицей $P(t_0)$ и не коррелирует с $w(t_k)$ и $v(t_{k+1})$, при любых значениях переменной k .

Ниже представлен алгоритм разработанной квадратно-корневой робастной модификации адаптивного фильтра Ахлагхи – Чжоу – Хуанга, предполагающий использование QR разложения и разложения Холецкого.

Инициализация:

$$\hat{x}(t_0|t_0) = \bar{x}(t_0), \quad P(t_0|t_0) = P(t_0), \quad P(t_0|t_0) = S_P(t_0|t_0)S_P^T(t_0|t_0); \quad \hat{R}(t_1) = R_0, \quad \hat{R}(t_1) = S_R(t_1)S_R^T(t_1); \quad \hat{Q}(t_0) = Q_0, \\ \hat{Q}(t_0) = S_Q(t_0)S_Q^T(t_0); \quad \alpha = \alpha_0, \quad \sigma = \sigma_0.$$

Выполнять в цикле по $k = \overline{0, N-1}$.

Одношаговое предсказание:

$$\hat{x}(t_{k+1}|t_k) = F(t_k)\hat{x}(t_k|t_k) + \Psi(t_k)u(t_k); \\ T^P(t_{k+1}) \begin{bmatrix} (F(t_k) S_P(t_k|t_k))^T \\ (\Gamma(t_k) S_Q(t_k))^T \end{bmatrix} = \begin{bmatrix} S_P^T(t_{k+1}|t_k) \\ 0 \end{bmatrix}.$$

Фильтрация:

$$L(t_{k+1}) = \exp\left(\frac{\varepsilon^T(t_{k+1})(S_R(t_{k+1})S_R^T(t_{k+1}))^{-1}\varepsilon(t_{k+1})}{2\sigma^2}\right); \\ T^2(t_{k+1}) \begin{bmatrix} S_R^T(t_k) & 0 \\ \sqrt{L(t_{k+1})}(H(t_{k+1})S_P(t_{k+1}|t_k))^T & S_P^T(t_{k+1}|t_k) \end{bmatrix} = \\ = \begin{bmatrix} S_B^T(t_{k+1}) & K_S^T(t_{k+1}) \\ 0 & S_P^T(t_{k+1}|t_{k+1}) \end{bmatrix}; \\ d(t_{k+1}) = y(t_{k+1}) - H(t_{k+1})\hat{x}(t_{k+1}|t_{k+1}); \\ T^R(t_{k+1}) \begin{bmatrix} \sqrt{\alpha} S_R^T(t_k) \\ \sqrt{1-\alpha} d^T(t_k) \\ \sqrt{1-\alpha} (H(t_{k+1})S_P(t_{k+1}|t_k))^T \end{bmatrix} = \begin{bmatrix} S_R^T(t_{k+1}) \\ 0 \\ 0 \end{bmatrix}; \\ K(t_{k+1}) = \sqrt{L(t_{k+1})} K_S(t_{k+1}) S_B^{-1}(t_{k+1}); \\ G(t_{k+1}) = (\Gamma^T(t_{k+1})\Gamma(t_{k+1}))^{-1}\Gamma^T(t_{k+1}); \\ T^Q(t_{k+1}) \begin{bmatrix} \sqrt{\alpha} S_Q^T(t_k) \\ \sqrt{1-\alpha} (H(t_{k+1})K(t_{k+1})\varepsilon(t_{k+1}))^T \end{bmatrix} = \begin{bmatrix} S_Q^T(t_{k+1}) \\ 0 \end{bmatrix}; \\ \varepsilon(t_{k+1}) = y(t_{k+1}) - H(t_{k+1})\hat{x}(t_{k+1}|t_k); \\ \hat{x}(t_{k+1}|t_{k+1}) = \hat{x}(t_{k+1}|t_k) + \sqrt{L(t_{k+1})} K_S S_B^{-1}(t_{k+1})\varepsilon(t_{k+1}).$$

Конец цикла.

Проведем сравнительный анализ фильтров Ахлаг-Чжоу-Хуанга, Изанлу-Фейкуриана-Джазди-Симона и разработанного фильтра на примере прямолинейного равноускоренного движения материальной точки под влиянием внешней среды.

Примем, что первая компонента вектора состояния отвечает за положение объекта, вторая – за его скорость, третья – за ускорение. Измеряется скорость объекта с некоторой ошибкой.

Запишем дискретные уравнения состояния и наблюдения [4]:

$$x(t_{k+1}) = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} x(t_{k+1}) + w(t_{k+1}),$$

$$y(t_{k+1}) = [0 \quad 1 \quad 0]x(t_{k+1}) + v(t_{k+1}), \quad k = \overline{0, N-1}.$$

Следуя статье[5], примем, что $\Delta t = t_{k+1} = 0.1$, $R_0 = 0.01$; $Q_0 = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$; $\bar{x}(t_0) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$; $P(t_0) = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$.

Проанализируем две ситуации. Сначала будем обрабатывать данные без выбросов, но с неточно заданными ковариационными матрицами шумов, полагая что $\hat{R}(t_1) = 100R_0$, а $\hat{Q}(t_1) = Q_0/100$. Далее будем обрабатывать данные с одиночными выбросами. Для этого смоделируем выборку, считая, что коэффициент загрязнения выборки $= 0.1$, а дисперсию шума аномальных наблюдений $\hat{R}_A = 10000R_0$.

Для ослабления зависимости результатов оценивания от выборочных данных проведем $M = 100$ различных запусков системы. Сравнение фильтров будем производить по значениям ARMSE. Результаты представим в таблице 1.

Таблица 1 – Значения ARMSE

Фильтр	ARMSE	
	Без выбросов	С одиночными выбросами
Адаптивный фильтр Ахлаги – Чжоу – Хуанга	<u>0.0140</u> ($\alpha = 0.999$)	1.2270 ($\alpha = 0.65$)
Робастный фильтр Изанлу – Фейкуриана – Джазди – Саймона	0.0943 ($\sigma = 5$)	0.3030 ($\sigma = 5$)
Разработанный фильтр	0.0142 ($\alpha = 0.999, \sigma = 5$)	<u>0.2187</u> ($\alpha = 0.999, \sigma = 5$)

Анализ таблицы 1 позволяет сделать вывод о том, что в ситуации, когда проверялась адаптивность фильтра, разработанный фильтр показал результаты незначительно хуже фильтра Ахлаги – Чжоу – Хуанга и заметно лучше, чем фильтр Изанлу – Фейкуриана – Джазди – Саймона.

В случае наличия аномальных выбросов в измерениях разработанный фильтр превзошел все остальные.

Целесообразно проанализировать эффективность разработанного фильтра на других задачах, в том числе практического характера.

Литература:

1. Цыганова Ю.В., Куликова М.В. Об эффективных методах параметрической идентификации линейных дискретных стохастических систем // Автоматика и телемеханика. – 2012. – №6. – С. 34-51.
2. Akhlaghi S., Zhou N., Huang Z. Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation // IEEE Power and Energy Conference General Meeting, 16-20 jul. 2017: Proceedings. – Chicago, 2017 – P. 1 – 5.
3. Izanloo R., Fakoorian S.A., Yazdi H.S., Simon D. Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise // Annual Conference on Information Science and Systems (CISS), 16-18 mar. 2016: Proceedings. – Princeton, 2016. – P. 500–505.
4. Chen B. Liu X., Zhao H., Principe J. Maximum correntropy Kalman filter. // Automatica. – 2017. – Vol. 76. – P. 70–77.

5. Чубич В.М., Прокофьева А.Э. Сравнительный анализ некоторых робастных фильтров для нестационарных линейных дискретных систем // Вестник Иркутского государственного технического университета. – 2017. – Т. 21. – № 12. – С. 123–137.

О ВЛИЯНИИ ПОВЕДЕНИЯ ВОДИТЕЛЯ И ПОГОДНЫХ РИСКОВ НА СТРАХОВЫЕ СЛУЧАИ

К.Г. Мадияров

**Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, kuan.mad@mail.ru**

Научный руководитель: Пестунов А.И., к. физ.-мат. н., доцент

В сложных климатических условиях, встречающихся на территории Казахстана особую значимость, приобретает влияние поведения водителя на прогнозирование страховых случаев. В данной статье описана модель, которая позволяет строить связь между показателями плохих погодных условий и поведением водителя с целью прогнозирования страховых случаев.

In the difficult climatic conditions found on the territory of Kazakhstan, the influence of driver behavior on predicting insured events becomes particularly important. This article describes a model that allows us to build a relationship between indicators of bad weather conditions and driver behavior in order to predict insurance claims.

Страховые компании Казахстана в своей практике могут сталкиваться с необходимостью прогнозирования страховых случаев с учетом погодных рисков. В исследованиях авторов [3], [4] модели CNN позволяют оценивать поведение водителей. Плоткина А.А. [1] Маглинова Т.Г. и Шупило О.М [2] предлагают внедрение AI в страховые компании для повышения точности моделирования рисков.

Отсутствуют комбинированные методы (на основе актуарных расчетов и машинного обучения), позволяющие прогнозировать страховые случаи, подверженные влиянию плохих погодных условий и поведению водителей одновременно. Это связано с тем, что поведение водителей, погодные факторы и суммы выплаты имеют очень слабую корреляцию.

Для подробного рассмотрения метода мы можем использовать выборку исторических данных с $n=550$ с данными согласно таблице 1. Уровень суровости погодных условий условно разделено на 5 уровней: 1 – ясная погода, 2 – небольшой дождь или снег, 3 – умеренный дождь или снег, 4 – сильный туман, гололед, 5 – ливень, снег, густой туман.

Таблица 1 – Выборка исторических данных

Сумма страховой выплаты (тг.).	Скорость ветра (м/с)	Видимость (км)	Уровень суровости погодных условий	Скорость водителя (км/ч)	Уровень агрессивности вождения
1	2	3	4	5	6
8028	3	13	3	131	6
9620	1	9	3	106	6

Продолжение таблицы 1

1	2	3	4	5	6
...
6937	5	9	3	120	8

**Пояснения к историческим данным:*

Уровень агрессивности вождения определяется на основе данных телеметрии и оценивается по 10 бальной шкале в зависимости от стиля вождения. Чем выше, тем более склонен к ДТП.

На сегодняшний день не существуют методов прогнозирования суммы выплат с учетом погодных рисков и поведения водителей, что приводит к повышению рисков для страховых компаний и необходимость перестрахования, чтобы уменьшить потери и риски.

Для установления связей и корреляции между суммой страховых выплат, поведением водителя и погодных рисков необходимо связать вышеизложенные метрики и рассчитать *комплексный погодный риск и скорректированный индекс риска*.

С помощью скорости ветра, видимости уровня суровости погодных условий можно вывести комплексный погодный риск R_w по формуле (V_w - скорость ветра, W_s – уровень суровости погодных условий, D - видимость):

$$R_w = V_w \times \frac{D}{W_s} \quad (1)$$

С помощью скорости водителя и уровня агрессивности вождения и комплексного погодного риска можно вывести скорректированный индекс риска R_c по формуле (V_d - скорость водителя, A – уровень агрессивности вождения):

$$R_c = \frac{V_d \times A \times R_w}{100} \quad (2)$$

Рассчитанные данные внесены в таблицу 2.

Таблица 2 – Расчет комплексного погодного риска R_w и скорректированного индекса риска R_c

Сумма страховой выплаты (тг.).	V_w	D	W_s	V_d	A	R_w	R_c
6887	3	13	3	131	6	0,8	80
6874	1	9	3	106	6	0,4	53
...
6894	5	9	3	120	8	1,8	102

С помощью данных индексов появляется возможность построить корреляцию между суммой страховых выплат, погодных условий и поведения водителя.

Строятся гипотезы для оценки влияния погодных условий и поведения водителя на сумму страховых выплат.

- *Первая нулевая гипотеза:* R_w не оказывает существенного влияния на сумму выплат.

- *Первая гипотеза:* R_w оказывает существенное влияние на сумму выплат. Тем самым повышается риск возникновения страховых случаев при плохих погодных условиях.

- *Вторая нулевая гипотеза:* R_c не оказывает существенного влияния на сумму выплат.

- *Вторая гипотеза:* R_c оказывает существенное влияние на сумму выплат. Тем самым повышается риск возникновения страховых случаев при плохих погодных условиях и при ненадлежащем вождении.

Для проверки данных гипотез строится модель множественной линейной регрессии, формула (α – базовая сумма выплат страхователям, равная 6906,59, β_1 – коэффициент влияния для R_w и он равен: -129,03, т.е. если R_w увеличивается на одну единицу, то сумма выплаты уменьшается на 129,03, β_2 – коэффициент влияния для R_c и равен: 0,995, т.е. увеличение R_c добавляет 0,995 к сумме выплат, X_1 – комплексный погодный риск R_w , X_2 – скорректированный индекс риска R_c , ε – случайная ошибка):

$$Y = \alpha + \beta_1 \times X_1 + \beta_2 \times X_2 + \varepsilon \quad (3)$$

После применения данной модели итоговая прогнозируемая сумма выплат с учетом плохих погодных условий и поведения водителя отражена в таблице 3.

Таблица 3 – Итоговые данные с прогнозируемой суммой выплат

Прогнозируемая сумма страховой выплаты (тг.).	V_w	D	W_s	V_d	A	R_w	R_c
8028	3	13	3	131	6	0,8	80
9620	1	9	3	106	6	0,4	53
...
6937	5	9	3	120	8	1,8	102

Результаты. Первая и вторая нулевые гипотезы опровергаются, т.к. погодный риск и риск ненадлежащего вождения в совокупности с погодным риском оказывают влияние на прогнозируемую сумму выплат, которая в свою очередь меняется из-за агрессивного стиля вождения при плохих погодных условиях увеличиваются сумма выплат, а из-за плохих погодных условий уменьшается. Например, при $R_w = 2,5$ и $R_c = 75$, будет составлять приблизительно 6659. Далее возникает необходимость поиска решений внедрений штрафного индекса для неаккуратных водителей.

Заключение. Можно строить предиктивную модель, которая прогнозирует сумму страховых выплат с учетом различных факторов. Это может помочь страховым компаниям при прогнозировании страховых случаев с учетом поведения водителей и плохих погодных условий.

Литература:

1. *Плоткина А.А.* Использование методов машинного обучения в имущественном страховании // *Высокие цифровые технологии: от анализа данных к искусственному интеллекту.* – 2021. – С. 260-263.
2. *Маглинова Т.Г., Шупило О.М.* Внедрение искусственного интеллекта в страховую отрасль // *Международный журнал гуманитарных и естественных наук.* – 2022. – № 5-4 (68). – С. 142-145. – doi:10.24412/2500-1000-2022-5-4-142-145.
3. *Varaich Z.A., Khalid S.* Recognizing Actions of Distracted Drivers using Inception v3 and Xception Convolutional Neural Networks // *Proceedings of the 2019 International Conference on Advanced Computing and Communication Systems (ICACS).* – 2019. – P. 347-351. – doi: 10.23919/ICACS.2019.8689131.
4. *Li Y., Sun C., Hu Y.* Whale Optimization Algorithm-Based Deep Learning Model for Driver Identification in Intelligent Transport Systems // *Computers, Materials & Continua.* – 2023. – Vol. 77. – P. 98-106. – doi: 10.32604/cmc.2023.035878.

КЛИЕНТ СЕРВЕРНОЕ РЕШЕНИЕ ДЛЯ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ

М.И. Ефиркин, И.И. Дыминский

**Новосибирский государственный технический университет,
г. Новосибирск, onirs@corp.nstu.ru**

Научный руководитель: Якименко А.А., к.т.н., доцент

Представлен прототип клиент-серверного решения для анализа временных рядов. Прототип разработан для обработки и визуализации данных о сейсмических событиях. Решение позволяет рассчитывать и отображать ключевые метрики событий, а также предоставляет доступ к самим данным в структурированном формате. Пользовательский интерфейс решения дополнительно предоставляет возможность фильтрации данных, что упрощает обработку больших массивов информации и обеспечивает интуитивный доступ к ключевым показателям.

The paper presents a prototype of a client-server solution for time-series analysis. This prototype is designed to process and visualize earthquake data. The solution provides calculation and display of key metrics, as well as presentation of the data itself in a structured format. The system's user interface provides data visualization options and data filters, which simplifies the processing of large data sets and provides intuitive access to key metrics.

Введение

В современном мире объемы собираемых данных постоянно растут, охватывая все сферы человеческой деятельности. Одной из наиболее распространенных форм представления таких данных являются временные ряды — последовательности упорядоченных во времени числовых показателей. Анализ временных рядов позволяет выявлять тенденции, прогнозировать будущее поведение систем и принимать обоснованные решения [1]. Однако обработка временных рядов требует специализированных инструментов, которые должны не только предоставлять возможности для анализа данных, но и быть интуитивно понятными и удобными в использовании для пользователя.

Существующие решения для анализа временных рядов часто сталкиваются с рядом проблем, касающихся удобства интерфейса и доступности функционала [2]. Отсутствие интуитивно понятного интерфейса и гибкости в настройке отображения данных часто затрудняют процесс анализа [3-4]. Из-за чего сильно уменьшается эффективность работы специалистов, работа которых связана с анализом временных рядов.

Для преодоления вышеописанных ограничений необходимо разрабатывать интерфейсы и инструменты, способные максимально

гибко подстраиваться под потребности пользователей. В данной работе представлен прототип одного из таких программных решений.

Описание архитектурного решения

Предложенное решение построено на базе клиент-серверной архитектуры, где серверная (бэкенд) и клиентская (фронтенд) части четко разделены и могут быть развёрнуты на разных вычислительных машинах, что обеспечивает высокую производительность и гибкость системы. Такая структура позволяет легко масштабировать приложение в зависимости от увеличения объема данных или добавления новых функций, что особенно важно для аналитических систем. В отличие от монолитных архитектур, здесь клиент и сервер могут масштабироваться независимо, что упрощает внесение изменений в одну из частей приложения без влияния на работу остальных частей [5].

Структурная схема архитектуры предложенного решения, иллюстрирующая путь преобразования временных рядов между бэкендом и фронтендом представлена на рисунке 1.



Рис. 1 – схема взаимодействия прототипа клиент-серверного решения

Серверная часть системы отвечает за обработку данных, собранных в формат Parquet, который оптимизирован для работы с большими объемами информации. С помощью библиотеки pandas сервер сначала преобразует данные во временные последовательности, а затем рассчитывает по ним основные характеристики (метрики), такие как время прихода максимума, средний доверительный интервал и вольт-временная характеристика. Эти метрики могут быть получены на клиентской стороне.

Пользователям предоставляются два основных способа работы с данными. Первый способ — это программный доступ к рассчитанным метрикам через REST API, что позволяет интегрировать данные в модели машинного обучения и другие автоматизированные аналитические процессы. API предоставляет гибкость, делая данные

доступными для внешних приложений, и позволяет строить прогнозы или классификационные модели на основе временных рядов.

Второй способ предполагает визуальный анализ данных средствами клиентской части решения. Клиентское приложение позволяет отображать рассчитанные метрики в виде графиков и таблиц, а также предоставляет пользователю инструменты для настройки и фильтрации данных (рисунки 2-3). Это позволяет пользователям отслеживать тенденции и выявлять закономерности визуально. Также предусмотрена возможность выгрузки таблиц данных в формате CSV, а графиков – в PNG для использования в сторонних аналитических инструментах.

Одним из важных преимуществ данной архитектуры является поддержка многопользовательского режима. Благодаря серверной обработке, каждый пользователь получает возможность гибкой настройки фильтров и вариантов отображения временных рядов в рамках отдельной сессии без необходимости выполнения сложных расчетов непосредственно на клиентской стороне.

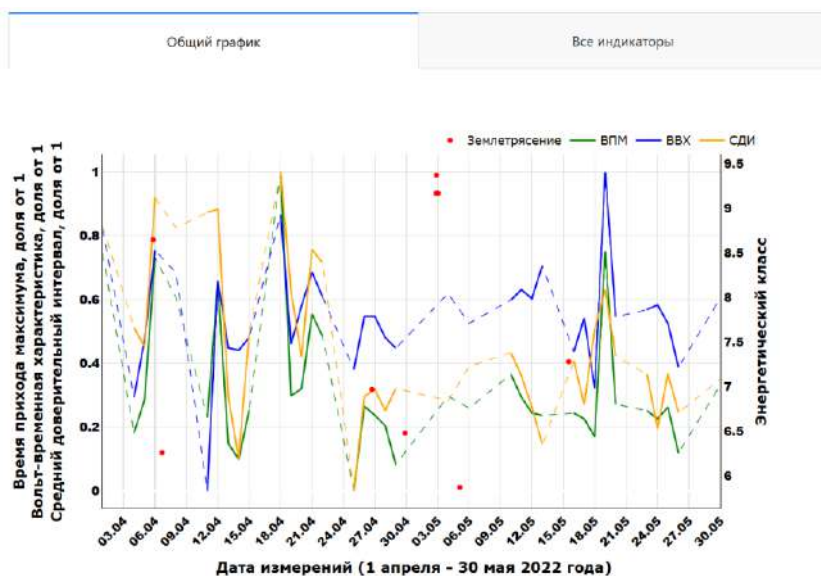


Рис. 2 – интерфейс визуализации временных рядов

Пункт наблюдения

Пункт наблюдения

Кегеты

Временной диапазон

01/04/2022 → 30/05/2022

Метрики

☒ Время прихода максимума

☒ Вольт-временная характеристика

☒ Средний доверительный интервал

☒ Нормировка метрик

☐ Показать количество землетрясений

Каталог сейсмических событий

Каталог

KNET

Радиус вокруг пункта (км) 50

Энергетический класс землетрясения (нижний предел)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Построить график

Рис. 3 – инструменты для настройки и фильтрации данных

Результаты

Основным результатом данной работы является прототип клиент-серверного решения, который предоставляет следующий функционал для работы с данными о сейсмических событиях:

- Выбор каталога сейсмических событий
- Выбор пункта наблюдения
- Ограничение данных по временному диапазону
- Радиус обнаружения землетрясений вокруг пункта наблюдения
- Ограничение по энергетическому классу землетрясений
- Просмотр и скачивание таблицы и графиков отфильтрованных данных
- Выборочный вывод нужных метрик
- Нормировка метрик
-

Заключение

Разделение задач между серверной и клиентской частью оптимизирует вычислительные процессы и создает основу для дальнейшего масштабирования системы. Гибкость серверной части позволяет легко интегрировать модели машинного обучения (такие как модели классификации или прогнозирования интенсивности сейсмических событий) расширяя аналитический функционал. Клиентская часть также может быть расширена путём добавления интерактивной карты сейсмических событий. Будущие модификации способны превратить представленное решение в полноценный инструмент для комплексного анализа временных рядов, позволяя

сопоставлять данные событий с временными рядами мониторинга и делать содержательные выводы. Система, способная к масштабированию и адаптации, будет полезна для выявления закономерностей и прогнозирования сейсмической активности.

Литература:

1. *Афанасьев В.Н., Юзбашев М.М.* Анализ временных рядов и прогнозирование: Учебник. – М.: Финансы и статистика, 2001. – 228 с.
2. *Иванова Е.В., Цымлер М.Л.* Обзор современных систем обработки временных рядов // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. – 2020. – №4. – URL: <https://cyberleninka.ru/article/n/obzor-sovremennyh-sistem-obrabotki-vremennyyh-ryadov> (дата обращения: 02.11.2024).
3. *Омар Абылайхан.* Методы создания качественного дизайна пользовательского интерфейса // Вестник науки. 2019. №10 (19). – URL: <https://cyberleninka.ru/article/n/metody-sozdaniya-kachestvennogo-dizayna-polzovatel'skogo-interfeysa> (дата обращения: 02.11.2024).
4. *Назаренко Н.А., Падерно П.И.* Влияние интерфейса на состояние и здоровье оператора // Биотехносфера. – 2009. – №6. – URL: <https://cyberleninka.ru/article/n/vliyanie-interfeysa-na-sostoyanie-i-zdorovie-operatora> (дата обращения: 02.11.2024).
5. *Сабиров Д.А.* Микросервисная архитектура на frontend // Научный журнал. – 2021. – №7 (62). – URL: <https://cyberleninka.ru/article/n/mikroservisnaya-arhitektura-na-frontend> (дата обращения: 02.11.2024).

АВТОМАТИЗАЦИЯ ПРОЦЕССА ПОСТРОЕНИЯ ДАШБОРДОВ В СИСТЕМЕ БИЗНЕС-АНАЛИТИКИ УНИВЕРСИТЕТА

К.А. Левковский

Новосибирский государственный технический университет,

г. Новосибирск, levkovskij.2019@stud.nstu.ru

Научный руководитель: Стасышин В.М., к.т.н., доцент

В данной работе описывается программное решение, описывающее оптимизацию построения дашбордов в Apache Superset с использованием шаблонизатора Handlebars для проректоров и начальников управлений Новосибирского государственного технического университета.

This paper describes a software solution describing the optimization of dashboard construction in Apache Superset using Handlebars templating engine for vice-rectors and heads of departments of Novosibirsk State Technical University.

Apache Superset — это мощная платформа для визуализации данных и бизнес-аналитики, которая позволяет пользователям создавать интерактивные дашборды и отчеты на основе различных источников данных. Эта платформа предлагает широкий набор инструментов для работы с данными, включая графики, таблицы, карты, фильтры и многое другое. Superset предоставляет возможность интеграции с различными базами данных и системами хранения данных, что делает её универсальным инструментом для анализа и визуализации данных.

На данный момент в Новосибирском государственном техническом университете (НГТУ) в Apache Superset развёрнуто более 100 дашбордов. Основными пользователями этих дашбордов являются проректоры и начальники управлений. В своей работе они используют одновременно более 10 дашбордов с разными данными, и на постоянное переключение между ними уходит много времени. Это создаёт неудобства и снижает эффективность работы, так как пользователи вынуждены тратить время на поиск и открытие нужных дашбордов.

Для решения этой проблемы был разработан скрипт на языке Python для отображения необходимых графиков в Apache Superset с использованием шаблонизатора, написанного на JavaScript – Handlebars. Этот скрипт позволяет автоматизировать процесс отображения ключевых метрик и графиков, что значительно упрощает работу пользователей и повышает их продуктивность.

Рассмотрим этапы реализации скрипта направленного на оптимизацию просмотра необходимых графиков:

1. Подготовка данных:
 - Сбор данных: на первом этапе нам необходимо собрать все данные, которые будут отображаться на дашборде.
 - Предобработка данных: предобработка данных для обеспечения их корректного отображения. Это включает в себя удаление дубликатов, заполнение пропущенных значений, нормализацию данных и другие операции.
2. Настройка Apache Superset:
 - Создание подключений: настройка подключения к источникам данных в Apache Superset. Это включает в себя настройку подключений к базам данных, API.
 - Создание датасетов: создаются датасеты в Apache Superset, которые будут использоваться для построения графиков и метрик. Это делается через API Apache Superset.
3. Разработка шаблонов:
 - Создание шаблонов Handlebars: разработка шаблонов для отображения графиков и метрик с использованием шаблонизатора

Handlebars. Шаблоны должны быть гибкими и позволять динамически изменять данные и параметры отображения.

- Интеграция шаблонов с данными: интеграция шаблонов с данными, которые были собраны и предобработаны на первом этапе. Это позволит автоматически генерировать графики и метрики на основе текущих данных.

4. Разработка скрипта на Python:

- Чтение данных: написан код на языке Python для чтения данных из источников и их предобработки.

- Генерация графиков: используя шаблоны Handlebars, генерируем графики и метрики на основе данных.

- Интеграция с Apache Superset: интеграция скрипта с Apache Superset для автоматического отображения графиков и метрик на дашбордах.

5. Тестирование и отладка:

- Тестирование: провести тестирование скрипта на различных наборах данных и в различных сценариях использования. Необходимо убедиться, что графики и метрики отображаются корректно и соответствуют ожиданиям пользователей.

- Отладка: в случае обнаружения ошибок или недочетов, провести отладку скрипта и внести необходимые изменения.

6. Развертывание и документирование:

- Развертывание: необходимо развернуть скрипт в рабочей среде и настроить его автоматическое выполнение по расписанию или при изменении данных.

- Документирование: подготовка документации по использованию скрипта и шаблонов. Включение в документацию описание всех этапов реализации, примеры использования и рекомендации по настройке.

Литература:

1. *Краснощечков Д.С.* Open - source решения в современной сфере business intelligence // Потенциал устойчивого инновационного развития: концепции, модели и практическое приложение. – СПб.: Общество с ограниченной ответственностью «Агентство международных исследований» (Уфа), 2023. – С. 91-93.

2. *Akbar R., Silvana M., Hersyah M.H. and Jannah M.* Implementation of Business Intelligence for Sales Data Management Using Interactive Dashboard Visualization in XYZ Stores // 2020 International Conference on Information Technology Systems and Innovation ICITSI 2020 – Proceedings, 2020. – P. 242-249

3. *Guilherme H. Soares, Miguel A. Brito.* Business Intelligence Over and Above Apache Superset, 2023 18th Iberian Conference on Information Systems and Technologies (CISTI).

4. API // Apache Superset – URL: <https://superset.apache.org/docs/api/> (дата обращения: 30.10.2024).

5. API // Handlebars. – URL: <https://handlebarsjs.com/api-reference/> (дата обращения: 30.10.2024).

ИСПОЛЬЗОВАНИЕ ШУМОПОДАВЛЕНИЯ ПОСЛЕ ПРИМЕНЕНИЯ BSS ДЛЯ ПОВЫШЕНИЯ КАЧЕСТВА РАСПОЗНАВАНИЯ РЕЧИ ПРИ НАЛИЧИИ ИСТОЧНИКОВ ШУМА

А.А. Пятанин

**Новосибирский государственный технический университет,
г. Новосибирск, alexander.pyatanin@gmail.com**

Научный руководитель: к.т.н., доцент каф. ПМт И.М. Ступаков

В данной статье рассматривается применение методов шумоподавления для улучшения качества распознавания речи в условиях шума. Исследование включает анализ методов Spectral Subtraction, Iterative Wiener Filtering, Noisereduce и RNNoise wrapper. Шумоподавление применяется после использования методов AuxIVA и WPE. Комплексный подход к шумоподавлению показывает значительное улучшение качества распознавания речи.

This paper explores the application of noise reduction methods to improve the quality of speech recognition in noisy environments. The study includes an analysis of Spectral Subtraction, Iterative Wiener Filtering, Noisereduce, and RNNoise wrapper methods. Noise reduction is applied after using AuxIVA and WPE methods. A comprehensive approach to noise reduction demonstrates significant improvements in speech recognition quality.

В данной работе будет использоваться online подход с комбинацией методов AuxIVA и WPE, предложенный в работе [1] в 2024 г.

AuxIVA (Auxiliary Function-based Independent Vector Analysis) – это метод слепого разделения источников (BSS), который используется для разделения смешанных аудиосигналов на их первоначальные компоненты. AuxIVA, в отличие от IVA, использует вспомогательные функции для упрощения оптимизации и улучшения сходимости.

WPE (Weighted Prediction Error) – алгоритм дереверберации, который используется для уменьшения реверберации в аудиосигнале. Реверберация – это эффект, возникающий в результате множественных

отражений звука от поверхностей в помещении, который может ухудшить качество аудио и сделать речь менее понятной.

Для определения эффективности работы алгоритма будем использовать WER. Word Error Rate – это метрика, используемая для оценки точности систем распознавания речи. Она измеряет количество ошибок, которые система делает при преобразовании речи в текст.

Как модель распознавания текста будем использовать vosk-model-small-ru-0.22 от проекта с открытым исходным кодом VOSK [4].

Для определения методов, которые стоит реализовать, воспользуемся некоторыми готовыми решениями и проверим их работоспособность. Тестирование проводилось на одном наборе данных с изменением основных параметров:

1. Single Channel Denoising

Pyroomacoustics предлагает три метода для подавления шума в одноканальных записях: Subspace Approach, Iterative Wiener Filtering и Spectral Subtraction. Subspace Approach был исключен из тестирования из-за высоких временных затрат.

Iterative Wiener Filtering (IWF) – это метод шумоподавления, который применяет фильтр Винера к каждому кадру сигнала и обновляет его на каждой итерации. Параметры включают: длину кадра (`frame_len`), количество коэффициентов LPC (`lpc_order`), количество итераций (`iterations`), коэффициент сглаживания (`alpha`) и порог для различия между кадрами (`thresh`).

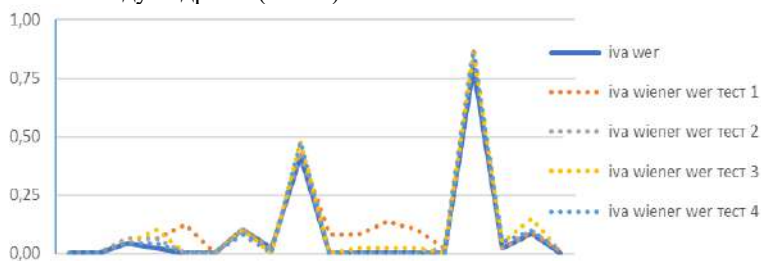


Рис. 1 Результаты применения IWF после AuxIVA и WPE

Spectral Subtraction (SS) – это метод шумоподавления, который вычитает оцененный спектр шума из спектра сигнала в частотной области [2]. Параметры включают: размер FFT (`nfft`), максимальное уменьшение в дБ (`db_reduc`), количество кадров для оценки шума (`lookback`), фактор переоценки (`beta`) и экспонентный фактор (`alpha`).

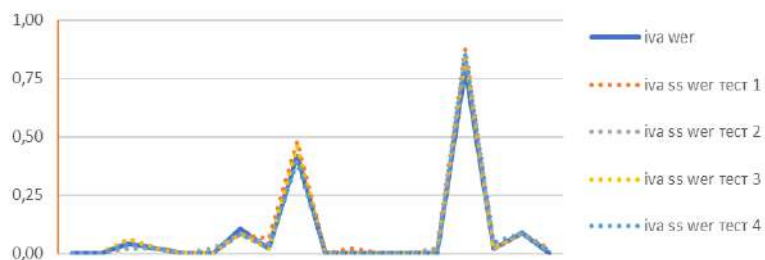


Рис. 2 Результаты применения SS после AuxIVA и WPE

2. Noisereduce

Noisereduce – библиотека для снижения шума в Python, основанный на методе спектрального затворения. Он включает два алгоритма:

Стационарное снижение шума (SRN): Параметры включают количество стандартных отклонений выше среднего значения для установки порога между сигналом и шумом (`n_std_thresh_stationary`).

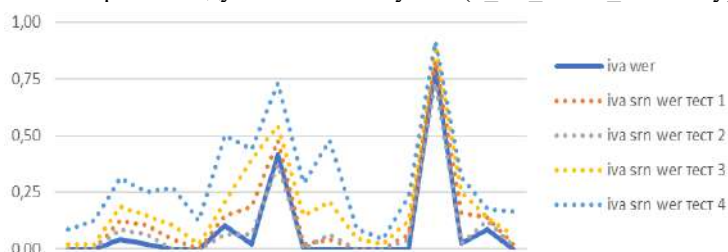


Рис. 3 Результаты применения SRN после AuxIVA и WPE

Нестационарное снижение шума (NRN): Параметры включают множитель, применяемый к пороговому значению для определения шума (`thresh_n_mult_nonstationary`).



Рис. 4 Результаты применения NRN после AuxIVA и WPE

Но для NRN требуется знать соотношение сигнала к шуму.

RNNNoise wrapper – это обёртка на Python, которая позволяет использовать RNNNoise. RNNNoise – это рекуррентная нейронная сеть с ячейками GRU, предназначенная для подавления шума в аудиозаписях.

Исследовано две версии модели RNNNoise: базовая и русская. RNNNoise (ru) не показал значительных результатов, а RNNNoise (basic) находила множество звуков из других языков.

По итогам замеров было решено реализовать собственный метод Spectral Subtraction, внедрить его в решение и подобрать оптимальные параметры для его использования. Параметры в полученной реализации: `past_frames` – это количество предыдущих фреймов, используемых для подавления шума; `denoise_db_reduction` – это максимальное значение подавления шума в децибелах для каждого бина спектра; `denoise_beta` – это параметр, который контролирует степень подавления шума в низкочастотной области; `denoise_alpha` – это параметр, который контролирует степень подавления шума в высокочастотной области.

1. Для подбора оптимальных параметров серия тестов была проведена следующим образом:

2. Создадим малую выборку тестирования для определения наиболее стабильных значений хотя бы одного параметра и зададим набор для перебора;

3. Далее зафиксируем наиболее стабильный параметр, увеличим выборку и зададим новый набор для перебора;

4. Будем сужать разброс значений и увеличивать выборку до тех пор, пока не получим определенное значение для каждого параметра.

Получив таким образом оптимальные параметры, убедимся, что реализованное шумоподавление релевантно. Для этого протестируем работу алгоритма на нескольких наборах данных, содержащих зашумленные звуковые данные на русском языке.

Таблица 1 – Эффективность распознавания речи

Набор данных	Средний WER оригинальной записи, %	Средний WER используя WPE и AuxIVA, %	Средний WER используя WPE, AuxIVA и SS, %
17.01.24	40,27	7,30	6,78
08.05.24	32,34	9,76	10,08
13.05.24	53,13	20,86	11,48

Каждый набор данных имеет свои особенности: набор от 17.01.24 содержит четкую и структурированную речь, загрязненную бытовыми

шумами или музыкой; набор от 08.05.24 ориентирован на тестирование отраженной речи, что может приводить к тому, что полезная информация воспринимается как шумы; набор от 13.05.24 наиболее крупный и содержит различные шумовые ситуации, такие как подкасты, фоновые диалоги, музыка со словами и т.д.

Методы шумоподавления в комплексе с методами слепого разделения сигналов значительно улучшают системы распознавания речи. Шумоподавление при определенных потерях зашумления дают значительную прибавку к производительности ASR. Для большей эффективности стоит регулировать параметры шумоподавления под определенное место использования.

Литература:

1. Ueda, T., Nakatani, T., Ikeshita, R., Kinoshita, K., Araki, S. and Makino, S. Blind and Spatially-Regularized Online Joint Optimization of Source Separation, Dereverberation, and Noise Reduction. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. – 2024.
2. Boll S.F. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. – 1979. – Vol. 27(2). – P. 113-120.
3. Jensen, J., & Hansen, J.H.L. Noise reduction in speech enhancement. *IEEE Signal Processing Magazine*. – 2015. – Vol. 32(2). – P. 42-51.
4. Kaldi, "Vosk," [Electronic resources]. – URL: <https://alphacephei.com/vosk/> [Accessed: 17.07.24].
5. Pyroomacoustics. Pyroomacoustics: A Python package for audio room simulation and array processing. – 2021. – URL: <https://pyroomacoustics.readthedocs.io/>. [Accessed: 17.07.24].

ПРОГРАММНЫЙ ИНСТРУМЕНТАРИЙ ДЛЯ СИСТЕМАТИЗАЦИИ РАБОТЫ ПОЛЬЗОВАТЕЛЕЙ СУПЕРКОМПЬЮТЕРОВ

В.С. Бурлаков, М.А. Ишутин

Новосибирский государственный технический университет,

г. Новосибирск, burlakov.2019@stud.nstu.ru

Научный руководитель: Городничев М.А.

Предлагаемая концепция персонального инструментария для управления вычислениями на основе формального описания предметной области позволяет автоматизировать создание и выполнение сценариев для локальных и удаленных вычислений. Подход упорядочивает работу с данными, обеспечивает

воспроизводимость экспериментов и отслеживание состояния расчетов, упрощая организацию вычислений на суперкомпьютерах.

The proposed concept of a personal toolkit for computation management, based on a formal domain description, enables the automation of creating and executing scripts for both local and remote computations, simplifying the process for researchers. This approach organises data handling, ensures experiment reproducibility, and tracks the status of calculations, reducing the complexity of human-supercomputer interaction.

Сложность проведения расчетов на суперкомпьютерах остается сдерживающим фактором для многих исследователей. В различных прикладных областях создаются специализированные приложения для высокоуровневой организации вычислительных экспериментов [1, 2], также известны проекты по созданию системных сервисов для управления данными экспериментов и удаленными вычислениями [3, 4, 5].

Цель работы – создание персонального программного инструментария для пользователей суперкомпьютерных центров (СКЦ), который позволит обеспечить автоматизацию проведения вычислительных экспериментов.

Предлагается концепция персонального программного инструментария пользователя для управления вычислениями, которая опирается на идеи, изложенные в работах [6, 7], и основана на формальном описании предметной области как множества величин и их функциональных зависимостей. Этот подход обеспечивает автоматизацию конструирования и выполнения вычислительных сценариев на персональном компьютере и удаленных вычислительных ресурсах, возможность отслеживания состояния расчетов, упорядочивание работы с данными, воспроизводимость компьютерных экспериментов.

Предлагается язык (основанный на TOML) для формального описания вычислительных экспериментов. Он включает в себя средства спецификации величин предметной области (переменных), операций, принимающих на вход и вычисляющих значения величин, и сценариев, составляемых из таких операций. Сценарий представляет собой двудольный ориентированный ациклический граф (DAG): одна доля вершин – это величины, другая – операции, которые их связывают. Для каждой операции пользователь указывает программу, реализующую соответствующую функциональную связь. Разработанный программный инструментарий обеспечивает выполнение таким образом специфицированных вычислительных сценариев, беря на себя реализацию рутинных действий, связанных с управлением

вычислениями и данными. В частности, обеспечивается передача программ и данных между персональным компьютером пользователя и суперкомпьютером, настройка окружения на суперкомпьютере, при этом учитываются особенности интерфейсов суперкомпьютерных систем.

Применение независимых окружений для изоляции каждой экспериментальной задачи снижает вероятность конфликтов и упрощает управление сложными экспериментами. Запуск вычислений через планировщик задач делает приложение масштабируемым, позволяя пользователям выполнять эксперименты на высокопроизводительном оборудовании СКЦ.

Рассмотрим схему работы инструментария. Для начала происходит проверка DAG на предмет корректности его структуры. После проверки происходит развертка DAG, что в текущей версии подразумевает линейное упорядочивание операций с учетом их зависимостей, но не исключает возможности параллельного выполнения независимых операций в новых версиях. На основе развернутого DAG создается план исполнения эксперимента, который также включает в себя задание значений входных величин, набор программ, реализующих операции, требования к аппаратному и программному обеспечению целевой вычислительной системы (ВС). Значения входных величин проходят проверку на соответствие ожидаемым типам данных. Заключительным этапом является запуск вычислительного эксперимента, в ходе которого осуществляется выбор ВС и отправка задач на выполнение через планировщик задач ВС.

Разработана модульная архитектура инструментария. Основой для пользовательского интерфейса (UI) является основной системный модуль, предоставляющий программный интерфейс (Application Programming Interface – API) другим компонентам. Это позволяет разрабатывать разные UI, такие как Command Line Interface (CLI) и Graphical User Interface (GUI). Основным модулем является Computer Workflow Application (CWA). Для хранения описания пользовательского вычислительного сценария используется модуль Computer Workflow (CW). Для выполнения описанного пользовательского сценария используется Operation Processor (OP), который в свою очередь обращается к модулям Data/Directory/Workload/Environment/Package manager, описывающим унифицированные интерфейсы взаимодействия с соответствующими инструментами и системой. В итоге, исполнение пользовательского сценария сводится к серии команд терминала, которые исполняются с помощью модуля Client, в том числе описывающего интерфейс подключения к удаленному оборудованию

(Рис. 1). На моменте прочтения производится валидация пользовательских описаний. Для удобства разработки и возможности расширения системы для некоторых модулей (Client, Environment, Package, Workload managers) предложен интерфейс взаимодействия модулей между собой.

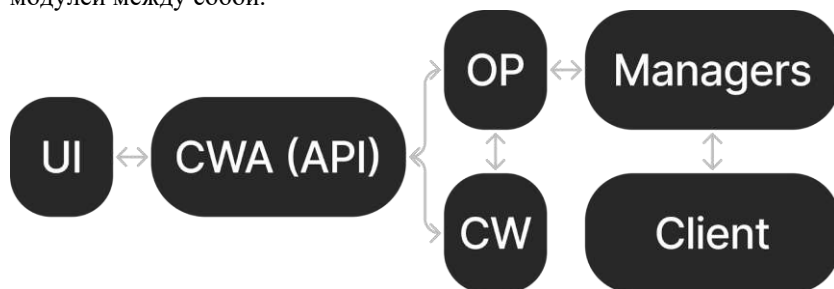


Рис. 1 – Общая архитектура инструментария

В качестве преимуществ реализации можно выделить: простоту архитектуры, унификацию описания окружения для развертывания программ, реализующих операции, и процедур запуска пользовательских экспериментов, расширяемость (возможность описания различных протоколов для клиентов, различных менеджеров, языков программирования для модулей), организация пользовательских данных, возможность запуска вычислений на удаленном оборудовании.

Использование описанной архитектуры и подхода к выполнению экспериментов позволяет пользователям суперкомпьютерных центров эффективно управлять экспериментами, снижая вероятность ошибок при выполнении рутинных действий и обеспечивая воспроизводимость результатов.

Литература:

1. *Chimakurthi, S.K., Reuss, S., Tooley, M. et al.* ANSYS Workbench System Coupling: a state-of-the-art computational framework for analyzing multiphysics problems // *Engineering with Computers*. – 2018. – Vol. 34. – P. 385-411. – doi: 10.1007/s00366-017-0548-4.
2. Petrel subsurface software [Электронный ресурс]. – URL: <https://www.slb.com/petrel> (дата обращения: 15.11.2024).
3. *Amstutz P. et al.* Common workflow language, v1. 0. – 2016.
4. *Jalili V. et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update // *Nucleic acids research*. – 2020. – Т. 48. – №. W1. – С. W395-W402.

5. *Singh P., Singh P.* Airflow // Learn PySpark: Build Python-based Machine Learning and Deep Learning Models. – 2019. – C. 67-84.
6. *Malyskin V.* Active knowledge, LuNA and literacy for oncoming centuries // Programming Languages with Applications to Biology and Security: Essays Dedicated to Pierpaolo Degano on the Occasion of His 65th Birthday. – 2015. – C. 292-303.
7. *Gorodnichev M., Lebedev D.* Semantic tools for development of high-level interactive applications for supercomputers // The Journal of Supercomputing. – 2021. – T. 77. – №. 10. – C. 11866-11880.

**ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ ВЛИЯНИЯ
ПРОСТРАНСТВЕННОГО РАСПОЛОЖЕНИЯ
ЦИЛИНДРИЧЕСКИХ СТОЛБИКОВ В МИКРОКАНАЛЕ
НА ГИДРОДИНАМИЧЕСКИЕ ХАРАКТЕРИСТИКИ**

**А.З. Булатова, Н.Б. Биккинина, О.А. Солнышкина
Уфимский университет науки и технологий,
г. Уфа, bulatova29@yandex.ru**

Ключевой задачей при проектировании микротеплообменников является повышение эффективности теплопередачи. Это достигается путем оптимизации скорости потока жидкости в каналах микротеплообменника для рассеивания большего количества тепла при меньших затратах. В этом исследовании изучаются различные конфигурации микроканалов с массивами цилиндрических столбиков. Для численного моделирования течения жидкости применяется ускоренный трехмерный метод граничных элементов (МГЭ). Изучается влияние расположения столбиков в микроканале на его пропускную способность и характер течения жидкости. Исследование демонстрирует связь между геометрией внутренней структуры микроканала и скоростью потока охлаждающей жидкости.

A key challenge in the design of micro heat sinks is to improve heat transfer efficiency. This is achieved by optimizing the fluid flow velocity in the micro heat sink channels to dissipate more heat at a lower cost. In this study, different microchannel configurations with cylindrical pin fin arrays are investigated. An accelerated three-dimensional boundary element method (BEM) is applied to numerically simulate the fluid flow. The influence of the pin fin arrays arrangement in the microchannel on its capacity and fluid flow pattern is studied. The study demonstrates the relationship between the geometry of the microchannel internal structure and the coolant flow velocity.

Современная электроника сталкивается с проблемой чрезмерного тепловыделения. Это связано с ростом мощности и миниатюризацией устройств, что приводит к повышению плотности мощности. Традиционные методы охлаждения становятся неэффективными, а их энергозатраты и влияние на окружающую среду возрастают. Перспективным решением является встраивание жидкостного охлаждения непосредственно в чип. Микроканальные радиаторы, используемые в таких системах, обладают высокой компактностью, то есть отношением площади удельной поверхности к общему объему, что

позволяет эффективно отводить тепло. Однако это также может вызывать увеличение сопротивления потоку жидкости и сложность изготовления. Правильно подобранные размеры, геометрия и распределение элементов внутри микроканала являются ключевыми в повышении эффективности отвода тепла и оптимальном расходе охлаждающей жидкости в микроканале.

Данное исследование направлено на трехмерное численное моделирование однофазного потока в микроканале со сложной внутренней структурой. Рассматриваются конфигурации микроканалов с распределенными в шахматном порядке цилиндрическими элементами диаметром $d = 50$ мкм и высотой $h = 50$ мкм. Рассматривается медленное течение вязкой несжимаемой жидкости, на участке микроканала задается постоянный перепад давления. Течение жидкости описывается уравнением Стокса без учета инерционных эффектов и уравнением неразрывности. Задаются граничные условия периодичности на входе и выходе микроканала, а также условие прилипания на поверхности столбиков и стенок микроканала. Задача решается ускоренным трехмерным методом граничных элементов (МГЭ), который является наиболее эффективным при решении трехмерных задач в областях со сложной геометрией [1, 2].

При сохранении общего количества столбиков $N = 88$ менялось расстояние между ними $\Delta = (20, 25, 30, 35)$, в следствие чего изменялась пустотность каналов. Общая пустотность ϕ сечения микроканала в плоскости xOy рассчитывалась как отношение суммы площадей поперечного сечения столбиков к площади поперечного сечения всего канала: $\phi_1=59.95\%$ при $\Delta = 20$, $\phi_2=65.11\%$ при $\Delta = 25\%$, $\phi_3=69.34\%$ при $\Delta = 30$, $\phi_4=72.84\%$ при $\Delta = 35$. При увеличении расстояния между столбиками так же возрастала площадь поверхности контакта. Проведены расчеты скорости потока и пропускной способности каналов при различных расстояниях между элементами. На рис. 1 показаны поля модуля скорости потока в микроканалах.

По рисункам видно распределение скорости в микроканале. Максимальные значения скорости потока достигаются в продольных участках между столбиками. При увеличении расстояния между цилиндрическими элементами скорость потока возрастает в следствие уменьшения гидравлического сопротивления за счет увеличения площади фильтрации жидкости.

Кроме картин течения в работе проанализирована пропускная способность микроканала в зависимости от расстояния между элементами. Получено, что при увеличении значения Δ улучшается пропускная способность рассматриваемых микроканалов. Данный

результат важен для дальнейшего развития исследования в области конструирования микротеплообменников, поскольку пропускная способность канала микротеплообменника напрямую влияет на эффективность отвода тепла в таком устройстве. Показано, что оптимальное расположение столбиков при: $\Delta = 35$ повышает пропускную способность микроканалов.

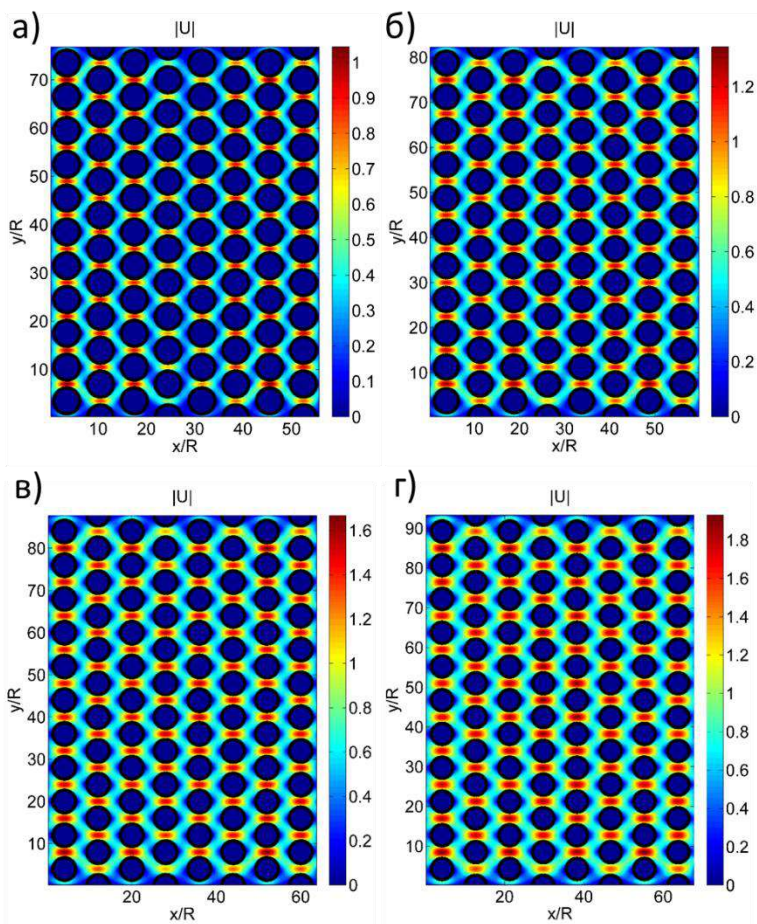


Рис.1 – поле скорости потока по модулю при: а) $\Delta = 20$, б) $\Delta = 25$, в) $\Delta = 30$, г) $\Delta = 35$

Известно, что увеличение площади контакта канала с теплоносителем может привести к повышению эффективности

теплоотвода. Важно найти баланс между пропускной способностью канала и площадью контакта, чтобы добиться максимального отвода тепла при минимальных затратах. Данное исследование предлагает гидродинамические основы для выбора оптимальной геометрии микроканалов. Следующим шагом является оценка теплопередачи для полной оптимизации конструкции микрофлюидных систем охлаждения. Это является перспективным направлением исследований с большим научным и практическим потенциалом.

Исследование выполнено при поддержке гранта Российского научного фонда № 24-19-00697 <https://rscf.ru/project/24-19-00697/>.

Литература:

1. *Абрамова О.А., Иткулова Ю.А., Гумеров Н.А., Ахатов И.Ш.* Трехмерное моделирование динамики деформируемых капель эмульсии методом граничных элементов и быстрым методом мультиполей на гетерогенных вычислительных системах. – Вычислительные методы и программирование, 2013. – Т. 14, № 4. – С. 438-450.
2. *Itkulova Y.A., Solnyshkina O.A., Gumerov N.A.* Toward large scale simulations of emulsion flows in microchannels using fast multipole and graphics processor accelerated boundary element method – ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE), Houston, TX. Vol. 7, PARTS A, B, C, D. – Houston, TX: American Society of Mechanical Engineers, 2012. – P. 873-881.

СРАВНЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ ПРЯМЫХ РЕШАТЕЛЕЙ РАЗРЕЖЕННЫХ СЛАУ НА CPU

Е.В. Васильев

**Новосибирский государственный технический университет,
г. Новосибирск, evgeniyvasiliev.01@gmail.com**

Научный руководитель: Домников П.А., к.т.н., доцент

В данной работе проведено сравнение вычислительной эффективности наиболее известных прямых решателей разреженных систем линейных алгебраических уравнений на центральном процессоре. Тестовые примеры взяты из практических задач вычислительной математики, эксперименты являются воспроизводимыми. Рассмотрены вопросы выбора и конфигурирования решателей.

The purpose of the study is to compare the performance of the most popular sparse linear systems direct solvers on CPU. The test samples belong to the practical

problems in computational mathematics and the tests are reproducible. Issues of choosing and configuring solvers are under consideration.

Рассмотрены последовательные и многопоточные версии решателей MKL PARDISO и SuperLU, последовательный решатель UMFPACK и параллельный решатель MUMPS на одном MPI-процессе. Эксперименты проведены при следующих условиях: процессор 11th Gen Intel Core i5-11400, 64 Гб RAM, ОС Windows 10, Intel MKL BLAS/LAPACK версии 2024.2, IEEE-арифметика двойной точности.

Описание тестовых задач приведено в таблице 1, все матрицы вещественнозначные. Матрицы 1-13 несимметричны и взяты из коллекции SuiteSparse [1]. Для них элементы правых частей сгенерированы с помощью равномерно распределённой на отрезке $[a, b]$ случайной величины, где a и b – модули минимального и максимального по модулю элементов матрицы; знак определялся с вероятностью $\frac{1}{2}$. Матрицы 14-23 – положительно полуопределённые, получены сложением вещественной и мнимой части СЛАУ гармонической задачи [2].

Таблица 1 – Описание тестовых задач

№	Название	n	nnz	condest	Предметная область
1	goodwin	7320	324784	7.58e+07	гидродинамика
2	raefsky4	19779	674195	1.78e+02	структурное модел.
3	wang3	26064	177168	4.99e+03	полупроводники
4	av41092	41092	1683902	3.56e+04	не указана
5	ckt11752_dc_1	49702	333029	1.19e+14	схемотехника
6	water_tank	60740	2035281	1.39e+06	гидродинамика
7	venkat01	62424	1717792	2.78e+03	гидродинамика
8	rajat23	110355	556938	8.65e+11	схемотехника
9	goodwin_127	178437	5778545	1.16e+05	гидродинамика
10	power197k	197156	3698357	3.08e+14	энергоснабжение
11	mac_econ_fwd500	206500	1273389	1.16e+13	экономика
12	torso3	259156	4429042	1,15e+02	электрофизика
13	marine1	400320	6226538	2,20e+07	океанология
14	3D1A_1	93961	2948497	4,27e+15	магнитотеллурическое зондирование; предобуславливатель для модели 3D1-A, описанной в [4]
15	3D1A_2	119729	3771857	6,29e+15	
16	3D1A_3	159267	5039163	6,99e+15	
17	3D1A_4	213433	6774193	3,69e+16	
18	3D1A_5	295882	9432922	1,09e+17	
19	DTM1_1	144318	4554510	2,95e+09	
20	DTM1_2	187622	5948342	3,36e+09	

21	DTM1_3	241752	7690632	2,75e+09	предобуславливатель для модели DTM1, описанной в [5]
22	DTM1_4	335170	1064379 4	3,56e+09	
23	DTM1_5	394326	1256669 4	3,68e+09	

В столбце `condest` таблицы 1 приведена оценка числа обусловленности снизу в 1-норме, полученная с помощью реализованной в SuperLU одной из разновидностей алгоритма 15.4 (LAPACK norm estimator) из [3]. Данный алгоритм является эвристическим и не предоставляет информации о точности получаемого решения. На практике оценка чаще всего не отклоняется от истинного значения больше, чем на порядок; тем не менее, существуют контрпримеры, для которых алгоритм даёт сколь угодно плохие оценки. Для полученных результатов оценка числа обусловленности коррелирует с величиной относительной невязки.

Основными факторами эффективности прямых решателей разреженных СЛАУ являются вычисление перестановки элементов матрицы для увеличения разреженности факторов и ресурсов параллелизма (NP-полная задача) и выбор метода обработки малых по модулю исключаяющих элементов. Для всех рассматриваемых решателей, кроме MKL PARDISO, характерно наличие алгоритмов замены исключаяющего элемента по строке, вызывающей изменение структуры разреженности факторов при перестановке столбцов, что может как повышать, так и понижать заполненность.

Результаты временных затрат приведены на рисунках 1 и 2. Мы стремились подбирать параметры решателей таким образом, чтобы при примерно одинаковом уровне невязки обеспечить наименьшие вычислительные затраты. Однако не для всех тестов получилось подобрать в достаточной степени удовлетворяющие этому требованию настройки. MKL PARDISO на плохо обусловленных задачах 14-18 показал худшую точность (см. таблицу 2 – решатели перечислены в том же порядке, что и на рисунке 1) в силу отсутствия замены исключаяющего элемента, несмотря на применение различных комбинаций критерия замены малых исключаяющих элементов на большие по модулю числа и количества шагов итерационной поправки. Многопоточный SuperLU показал замедление относительно последовательной версии на некоторых тестах из-за того, что значительная часть управления памятью и параллелизмом не автоматизирована и делегируется пользователю, что существенно затруднило выбор эффективной конфигурации.

При определении параметров решателей, реализующих замену ведущего элемента, применяются различные эвристики, например, выбор алгоритма переупорядочивания и принимаемой им структуры (A^T+A или A^TA), а также порога замены ведущего элемента, в зависимости от коэффициента симметрии матрицы. Подобные техники реализованы в пакетах UMFPACK и MUMPS, однако они нередко делают не лучший выбор, в связи с чем автоматизация подбора параметров представляется актуальным направлением исследований. Так, если параметры подбираются для класса задач со схожей структурой и обусловленностью матриц, можно применять методы оптимизации по параметрам для матриц небольшой размерности.

Согласно полученным результатам, при решении сравнительно хорошо обусловленных задач наиболее эффективен решатель MKL PARDISO, в то время как MUMPS несколько уступает по времени, но лидирует среди остальных решателей и является более устойчивым; оба пакета обладают привлекательным дополнительным функционалом (дополнение Шура, режим out-of-core и прочее). Наиболее надёжным алгоритмом переупорядочивания из использованных при проведении экспериментов (различные модификации алгоритма минимальной степени, реализующий алгоритм вложенных сечений METIS и PORD) является METIS (parMETIS), так как он не столь существенно уступает другим алгоритмам на небольших матрицах и выигрывает с увеличением размерности. Следует отметить довольно высокую точность решения СЛАУ прямыми методами: для многих задач уровень относительной невязки для различных алгоритмов факторизации отличается не более, чем на порядок, что свидетельствует о достижении некоторого уровня точности, определяемого обусловленностью задачи. Подчеркнём, что результаты исследования носят эмпирический характер, поскольку наилучшие конфигурации решателей нам неизвестны.

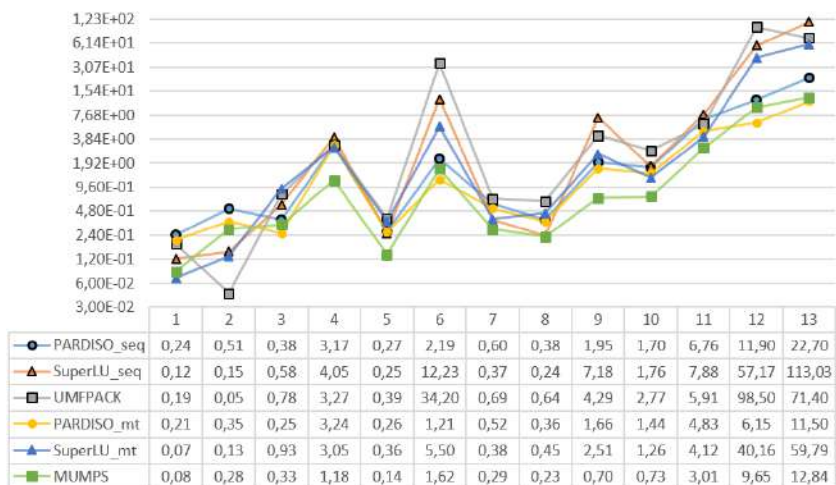


Рис.1 – Суммарное время решения задач 1-13 в секундах

Таблица 2 – Порядок относительной невязки

Таблица 12. Порядок отбора кандидатов																								
№		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Решатель	Ps	11	16	16	13	7	10	16	11	15	5	7	16	13	6	5	5	6	4	14	14	12	14	14
	Ss	11	16	15	14	7	9	16	11	14	4	7	16	12	13	13	13	13	16	14	14	13	13	13
	U	11	16	16	16	7	10	16	11	15	4	7	16	13	13	13	13	13	13	14	14	13	13	13
	Pm	11	16	16	13	7	10	16	11	15	4	7	16	13	7	5	6	6	4	14	14	12	14	14
	Sm	11	16	15	14	7	9	15	11	14	4	7	16	12	-	-	-	-	-	-	-	-	-	-
	M	11	16	16	16	7	10	16	11	15	4	7	16	13	10	13	13	12	12	14	14	12	13	13

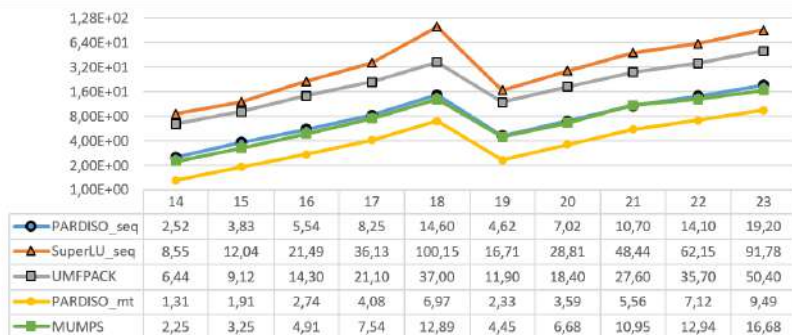


Рис.2 – Суммарное время решения задач 14-23 в секундах

Литература:

1. *Davis T.A., Hu Y.* The University of Florida sparse matrix collection. ACM Transactions on Mathematical Software (TOMS) 38 (1), 1
2. *Soloveichik Y.G., Persova M.G., Domnikov P.A., Koshkina Y.I. et al.* Iterative solver with folded preconditioner for finite element simulation of magnetotelluric fields // Computers and Geosciences. – 2022. – Vol. 169. – Art. 105244 (12p.). – doi 10.1016/j.cageo.2022.105244.
3. *Higham N.J.* Accuracy and Stability of Numerical Algorithms. 2nd edition. SIAM, Philadelphia, PA, 2002.
4. *Жданов М.С.* Методы моделирования электромагнитных полей (Материалы международного проекта COMMEMI) / М. С. Жданов, И. М. Варенцов, Н. Г. Голубев, В. А. Крылов. М.: Наука, 1990. – 198 с.
5. *Miensonopust, M.P., Queralt, P., Jones, A.G.* The 3D MT modellers. Magnetotelluric 3-D inversion — a review of two successful workshops on forward and inversion code testing and comparison. Geophys. J. Int. 193. – 2013. – P. 1216–1238.
6. *J. W. Demmel, S. C. Eisenstat, J. R. Gilbert, X. S. Li, and Joseph W. H. Liu.* A supernodal approach to sparse partial pivoting. SIAM J. Matrix Analysis and Applications. – 1999. – Vol. 20(3). – P. 720–755.
7. *T. A. Davis.* Algorithm 832: UMFPACK, an unsymmetric-pattern multifrontal method. ACM Trans. Math. Softw., 30(2):196–199, 2004.
8. *Amestoy P.R., Duff I.S., Koster J., L'Excellent J.-Y.* A fully asynchronous multifrontal solver using distributed dynamic scheduling. SIAM Journal on Matrix Analysis and Applications. – Vol. 23(1):15–41. – 2001.

РАСПАРАЛЛЕЛИВАНИЕ ПРОЦЕДУР МОДЕЛИРОВАНИЯ ХИМИЧЕСКИХ РЕАКЦИЙ ДЛЯ КОМПОЗИЦИОННОЙ МОДЕЛИ НЕФТИ

К.М. Виноградов

Новосибирский государственный технический университет,

г. Новосибирск, k.vinogradov.2019@stud.nstu.ru

Научный руководитель: Персова М.Г., д.т.н., профессор

В работе рассматривается проблема распараллеливания процедур моделирования химических реакций для композиционной модели нефти. Предлагается метод, который позволяет значительно ускорить вычисления за счёт использования параллельных вычислительных ресурсов. Результаты исследования демонстрируют повышение эффективности предложенного подхода при сохранении точности моделирования.

The paper considers the problem of parallelization of chemical reaction modeling procedures for a composite oil model. A method is proposed that allows to significantly enhance calculation performance through the use of parallel computing resources. The results of the study demonstrate an increase in the efficiency of the proposed approach while maintaining the accuracy of modeling.

В последние десятилетия моделирование углеводородных месторождений с учётом их многокомпозиционного состава приобрело особую важность. Это связано с необходимостью улучшения качества проектирования разработки и обустройства нефтяных и нефтегазоконденсатных коллекторов. Повышение точности описания фазового поведения таких систем позволит повысить качество управления разработкой и тем самым повысить общую экономическую эффективность разработки месторождения.

Многокомпозиционные модели нефти включают в себя большое число компонент, что значительно усложняет процесс моделирования химических реакций. Каждый компонент имеет свои уникальные свойства и правила взаимодействия с другими компонентами, что требует значительных вычислительных ресурсов для описания фазовых переходов в системе. Использование параллельных вычислительных ресурсов позволяет значительно ускорить процессы моделирования благодаря распределению задач между несколькими процессорами или ядрами при сохранении точности моделирования.

В качестве математической модели в данном исследовании используется закон Дарси, описывающий течение жидкостей и газов в пористой среде, и закон сохранения масс в пластовой системе. Математическая модель для расчёта давления имеет следующий вид:

$$-div(\sum_{m=1}^M \frac{\kappa^m}{\eta^m} \mathbf{K}(grad(P) + (0, 0, \rho^m g)^T)) = f, m = \overline{1, M}, \quad (1)$$

$$P|_{\Gamma_1} = P_{\Gamma_1}, \quad (2)$$

$$\sum_{m=1}^M \frac{\kappa^m}{\eta^m} \mathbf{K}(grad(P) + (0, 0, \rho^m g)^T) \Big|_{\Gamma_2} \cdot \vec{n} = \theta. \quad (3)$$

В соотношениях (1)-(3) P – давление, \mathbf{K} – тензор абсолютной проницаемости, g – ускорение свободного падения, M – число фаз, κ^m – относительная проницаемость m -й фазы, η^m – вязкость m -й фазы, ρ^m – плотность m -й фазы, Γ_1 – удалённая граница области, P_{Γ_1} –

давление на границе области, Γ_2 – границы скважин, θ – поток смеси, через скважины, f – источник, возникающий из-за изменения объёмов фаз.

Для решения задачи (1)-(3) используется метод конечных элементов [1]. На каждом временном шаге решается задача (1)-(3), что позволяет определить поле давления. С использованием поля давления вычисляются потоки фаз в расчётной области. Далее определяются объёмы фаз, перемещающиеся через границы ячеек конечно-элементной сетки, и обновляется распределение фаз и компонент в модели. Более детальное описание численного метода моделирования можно найти в работах [1-2].

Химические реакции между компонентами моделируются на основе равенства химических потенциалов [3]

$$y_i \varphi_i^{gas} = x_i \varphi_i^{liquid}, i = \overline{1, N_c}, \quad (4)$$

где x_i – молярная доля i -ой компоненты в фазе жидкости, y_i – молярная доля i -ой компоненты в фазе газа, φ_i^{liquid} – летучесть i -ой компоненты в фазе жидкости, φ_i^{gas} – летучесть i -ой компоненты в фазе газа, N_c – общее число компонент. При пересчёте давления равенство химических потенциалов может нарушиться. Отдельные компоненты могут переходить из жидкой фазы в газообразную, а другие могут конденсироваться из газа в жидкость. Чтобы определить молярные доли компонент на каждом элементе, которые будут соответствовать новому равенству химических потенциалов для нового распределения поля давления, необходимо решить систему уравнений [3]

$$\begin{cases} \ln(\varphi_i^{liquid}) - \ln(\varphi_i^{gas}) - \ln(K_i) = 0, i = \overline{1, N_c} \\ \sum_{i=1}^{N_c} \frac{(K_i - 1)z_i}{1 + (K_i - 1)\tilde{S}} = 0 \end{cases}, \quad (5)$$

где \tilde{S} – молярная доля газа в смеси на элементе, z_i – молярная доля компоненты в смеси, K_i – отношение y_i к x_i . Данная система уравнений является нелинейной, поскольку, согласно уравнению Пенга-Робинсона [3], логарифмы летучестей компонент и плотности фаз зависят от x_i , y_i . Для вычисления вязкости фаз использовалось уравнение Лоренца-Брея-Кларка [4]

$$[(\eta - \eta^*)\xi + 10^{-4}]^{1/4} = \alpha_4 \rho_r^4 + \alpha_3 \rho_r^3 + \alpha_2 \rho_r^2 + \alpha_1 \rho_r + \alpha_0, \quad (6)$$

где η^* – вязкость разреженного газа [3], коэффициент ξ вычисляется по формуле

$$\xi = \left(\sum_{i=1}^{N_c} x_i M_i \right)^{-1/2} \left(\sum_{i=1}^{N_c} x_i P_i^c \right)^{-2/3} \left(\sum_{i=1}^{N_c} x_i T_i^c \right)^{1/6}. \quad (7)$$

В соотношении (7) M_i – молярная масса i -ой компоненты, P_i^c , T_i^c – критические давление и температура i -ой компоненты, а значения α_j приведены в таблице 1.

Таблица 1 – Коэффициенты для формулы вязкости Лоренца-Брея-Кларка

α_0	α_1	α_2	α_3	α_4
0.0879560	-0.0217230	0.0664196	-0.0407580	0.0093324

Для распараллеливания процедур моделирования химических реакций был выбран стандарт OpenMP, так как он является одним из самых эффективных, известных и развитых инструментов, подходящих для этого. Его главным преимуществом является то, что он позволяет достичь ускорения с минимальными изменениями в существующих алгоритмах.

Для измерения ускорения проводились замеры только тех процедур, которые принимали участие в моделировании химических процессов. Вычисления проводились на шестиядерном процессоре с отключённой функцией гиперпоточности. Графики ускорения в зависимости от количества потоков представлены на рисунке 1.

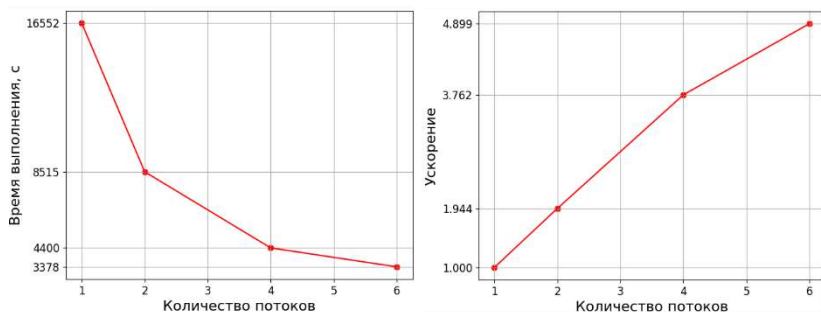


Рис.1 – Графики времени выполнения и ускорения от числа потоков.

Таким образом, предложенный метод распараллеливания процедур моделирования химических реакций показал свою эффективность. На исследуемых моделях было достигнуто ускорение примерно в 4.9 раз на 6 потоках по сравнению с выполнением на 1 потоке. Кроме того, проведенные исследования показали, что в полном цикле гидродинамического моделирования для композиционной модели нефти процедуры расчета химических реакций являются наиболее затратными и требуют порядка 30% общего вычислительного времени. Поэтому разработанный подход позволил значительно повысить эффективность гидродинамического моделирования в целом.

Литература:

1. Соловейчик Ю.Г., Рояк М.Э., Персова М.Г. Метод конечных элементов для решения скалярных и векторных задач // Новосибирск: НГТУ, 2007. – 896 с.
2. Soloveichik Y.G. [et al.] A method of FE modeling multiphase compressible flow in hydrocarbon reservoirs // Computer Methods in Applied Mechanics and Engineering. – 2022. – Vol. 390. – Art. 114468.
3. Лысов В.Г., Рыков Ю.Г. О вычислении фазового равновесия в задачах многокомпонентной фильтрации // Препринты ИПМ им. М.В. Келдыша. – 2014. – № 94. – 20 с. – URL: <http://library.keldysh.ru/preprint.asp?id=2014-94>.
4. Lohrenz, J., Bray, B.C., and Clark, C.R. "Calculating Viscosities of Reservoir Fluids from Their Compositions," JPT (October 1964), 1171.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ПАРАМЕТРОВ КАТУШКИ С ТОКОМ НА КАЧЕСТВО ПОЛЯ ЭЛЕКТРОМАГНИТА

Т.Р. Дербышева

**Новосибирский государственный технический университет,
г. Новосибирск, derbysheva.t@yandex.ru**

Научный руководитель: Рояк М.Э., д.т.н., профессор

В данной статье рассматривается проблема снижения однородности магнитного поля в магнитных детекторах в результате смещения катушки. Рассчитано распределение магнитной индукции B при различных смещениях катушки с помощью программного комплекса для численного моделирования Telma. Проведены исследования по подбору корректирующих токов, компенсирующих возникающую в результате смещения катушек неоднородность магнитного поля.

This paper considers the problem of magnetic field homogeneity reduction in magnetic detectors as a result of coil displacement. The distribution of magnetic induction B at different coil displacements is calculated using the numerical modeling software Telma. Correcting currents that compensate for the magnetic field inhomogeneity resulting from the coil displacement are selected.

Одним из ключевых факторов для обеспечения нормальной работы магнитного детектора является однородность магнитного поля, на которую влияют различные факторы, такие как качество изготовления катушки, ее положение внутри магнитопровода, а также различные сдвиги, возникающие в связи с большими размерами и массой детекторов. В связи с этим возникает необходимость вычислять однородность и исследовать возможности ее улучшения.

Рассмотрим вычисление однородности и ее изменение под влиянием некоторых из указанных факторов на примере детектора, имеющего высокие требования к однородности поля, модуль которого равен 0.5 Т.

Рассматриваемый магнит состоит из стального магнитопровода, а также трех сверхпроводящих катушек, расположенных вдоль продольной оси соленоида, и двух корректирующих медных катушек, размещенных на торцах. Сверхпроводящие катушки суммарно состоят из 1688 витков проводника с сечением 4.5 мм на 20 мм. Ток, протекающий в проводнике, равен 1790 А. Каждая из медных корректирующих катушек состоит из 34 витков проводника сечением 30 мм на 30 мм. Ток в проводнике составляет 3059 А. На рисунке (1) приведены графики распределения однородности для магнита в исходном состоянии.

Для рассматриваемого случая распределение однородности, показанное на рисунке 1 может быть улучшено при оптимизации токов в катушках. Для увеличения однородности было рассмотрено несколько вариантов изменения токов в корректирующих и сверхпроводящих катушках, которые представлены в таблице 1.

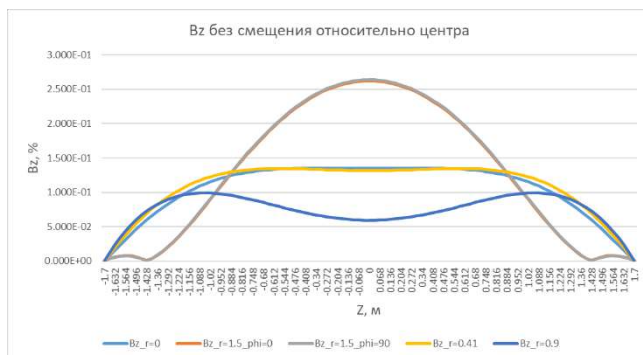


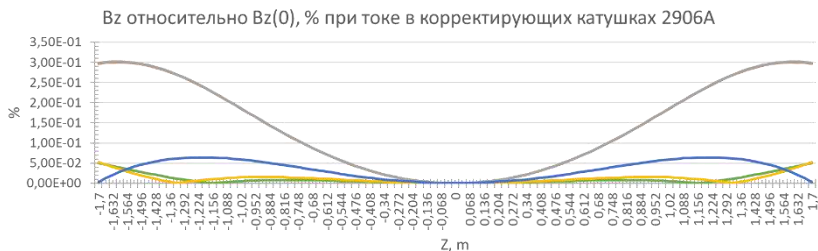
Рис. 1 – Однородность в % продольной компоненты B_z вектора магнитной индукции B вдоль оси Z на разных радиусах. Исходная модель

Таблица 1 – Варианты изменения тока для улучшения однородности

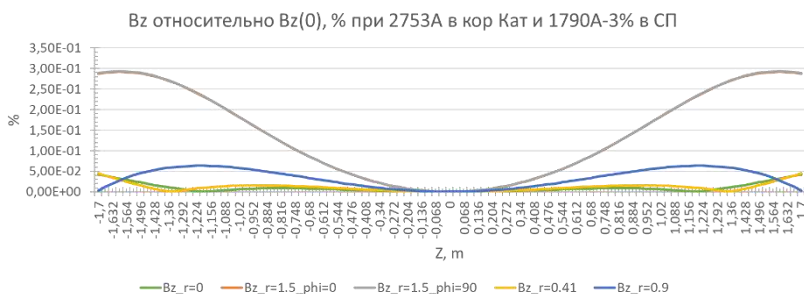
№	Ток в корректирующих катушках	Ток в сверхпроводящих катушках
1	2447 А	1790 А
2	2906 А	1790 А
3	2447 А	1781 А
4	2447 А	1772 А
5	2447 А	1799 А
6	2447 А	1807 А
7	2447 А	1754 А
8	2447 А	1700 А
9	2753 А	1790 А
10	2753 А	1736 А

Из представленных вариантов изменения были выбраны № 2 и № 10, при которых удалось получить максимальную однородность. Распределение однородности компоненты B_z вдоль Z для этих случаев представлено на рисунке 2.

Оценка влияния смещения катушек с идеальной геометрической оси на однородность поля является важным пунктом для понимания дальнейших свойств магнита в работе. Был рассмотрен вариант смещения параллельным переносом верхней катушки на 0.1 мм по X и на 2.47 мм по Y , а нижней на 0.42 мм по X и 1.2 мм по Y . На рисунке 3 показано распределение однородности компоненты B_z вдоль Z для рассматриваемого смещения.



а)



б)

Рис. 2 – Распределение однородности для изменения тока в катушках: а) № 2, б) № 10

При сравнении графиков на рисунках 1 и 3 видно, что максимальная разница в величине однородности составляет не более 0.05 %. Таким образом, рассмотренный вариант сдвига катушек не оказывает влияния на однородность поля детектора.

Исследование влияния сдвигов катушки на однородность магнитного поля в магнитных детекторах является важной задачей, необходимой для обеспечения точности измерений. Результаты данного исследования позволяют выявить факторы, которые могут приводить к искажению магнитного поля, и разработать методы корректировки для достижения наиболее однородного поля. В рассмотренном в данной работе магните показана возможность улучшать однородность с помощью изменения корректирующих токов.

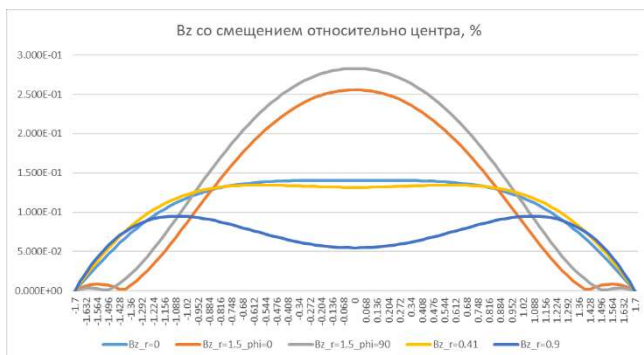


Рисунок 3 – Однородность в % продольной компоненты B_z вектора магнитной индукции B вдоль оси Z на разных радиусах. Модель со смещениями катушек

Литература:

1. Hartmann F., Sharma A. Multipurpose detectors for high energy physics, an introduction // Nuclear Instruments and Methods in Physics. – 2012. – Vol. 666. – P. 1-9. – <https://doi.org/10.1016/j.nima.2011.11.023>.
2. Golovatyuk, V., Kekelidze, V., Kolesnikov, V. et al. The Multi-Purpose Detector (MPD) of the collider experiment // The European Physical Journal A. – 2016. – Vol. 52. – P. 212. – <https://doi.org/10.1140/epja/i2016-16212-1>.
3. Zhu, X., Xing, M., Ye, J. et al. Design and optimization of a novel solenoid with high magnetic uniformity // Sci Rep. – 2024. – Vol. 14. – <https://doi.org/10.1038/s41598-024-76501-y>.
4. Royak M.E., Stupakov I.M. and Kondratyeva N.S. Coupled vector FEM and scalar BEM formulation for eddy current problems // 13th International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE), Novosibirsk, Russia. – 2016. – P. 330-335. – doi: 10.1109/APEIE.2016.7806481.

РАЗРАБОТКА И РЕАЛИЗАЦИЯ ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ И ИНТЕРПРЕТАЦИИ ДАННЫХ ИНДУКЦИОННОГО КАРОТАЖА

М.В. Иванов, М.Г. Персова

Новосибирский государственный технический университет,
г. Новосибирск, m.ivanov.2017@stud.nstu.ru

Описывается решение обратной задачи индукционного каротажа. Приведены результаты решения обратной краевой задачи.

The solution to an inverse problem of induction logging is described. The results of solving the inverse boundary value problem are presented.

Введение. Метод высокочастотного индукционного каротажного изопараметрического зондирования (ВИКИЗ) [1] – это один из методов скважинной геофизической разведки, в котором детально исследуется строение разреза скважины с помощью спуска-подъёма в неё геофизического каротажного зонда. Данный метод основан на измерении относительных фазовых характеристик высокочастотного электромагнитного поля, создаваемого генераторными катушками и принимаемого измерительными катушками. Метод ВИКИЗ имеет небольшой радиус исследования околоскважинного пространства (от нескольких сантиметров до нескольких метров), но обладает высокой детальностью, позволяющей не только определить с точностью до сантиметров глубину залегания пласта, но даже характер изменения самого пласта на всей его небольшой мощности. Преимуществом данного метода является высокое пространственное разрешение, повышающее эффективность исследования маломощных пластов.

Под интерпретацией данных понимается определение (восстановление) параметров среды по данным, полученным в результате поисково-разведочных работ. В данной работе задача интерпретации данных электромагнитного каротажа заключается в восстановлении по каротажным данным значений удельной электрической проводимости в зоне проникновения и в области за зоной проникновения и сводится к минимизации функционала квадрата отклонения практических и расчетных данных.

Математическая модель. Двумерная осесимметричная краевая задача, описывающая поведение гармонического по времени электромагнитного поля [2], определяется скалярным уравнением

$$-div(\frac{1}{\mu} grad A_{\varphi}) + (i\sigma\omega - \varepsilon\omega^2 + \frac{1}{\mu r^2})A_{\varphi} = J_{\varphi}$$

с краевыми условиями

$$A_{\varphi}|_{S_1} = 0, \frac{1}{\mu} \frac{\partial A_{\varphi}}{\partial n} \Big|_{S_2} = 0.$$

Решение обратной задачи. Решением обратной задачи является восстановленная геоэлектрическая модель среды по практическим данным. Искомые параметрами при решении обратной задачи являются:

- значения удельной электрической проводимости σ внутри блоков, ограниченных вертикальными и горизонтальными линиями;
- z-координаты границ между слоями;
- r-координаты радиальных границ внутри каждого слоя.

Решение обратной задачи основано на минимизации функционала:

$$\Phi(b) = \sum_{i=1}^N \left(V_i(b) - \tilde{V}_i \right)^2 + \sum_{m=1}^M a_m \left(\Delta b_m \right)^2.$$

Результаты. Для верификации разработанной программы решения обратной задачи использовались синтетические данные индукционного каротажа, полученные с помощью конечноэлементного моделирования данных для моделей сложнопостроенных коллекторов.

Результаты верификации подтвердили правильность реализации численных процедур.

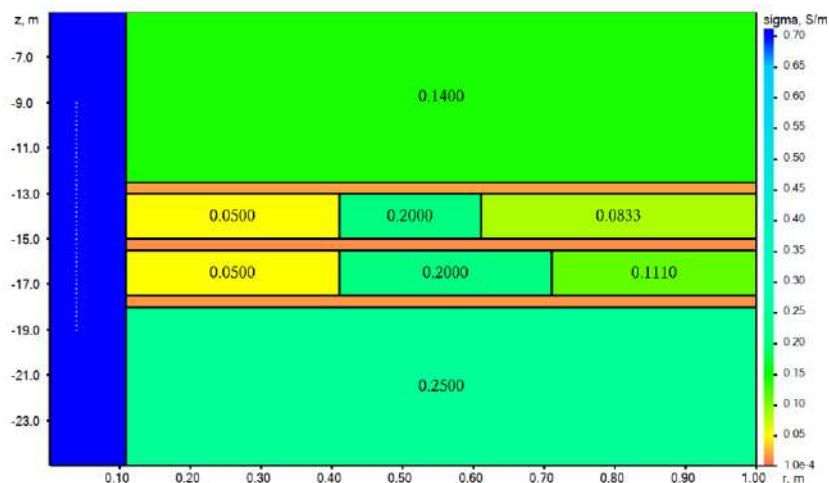


Рис. 1 – Истинная геоэлектрическая модель нефтеводонасыщенного коллектора, осложнённого карбонатными прослоями.

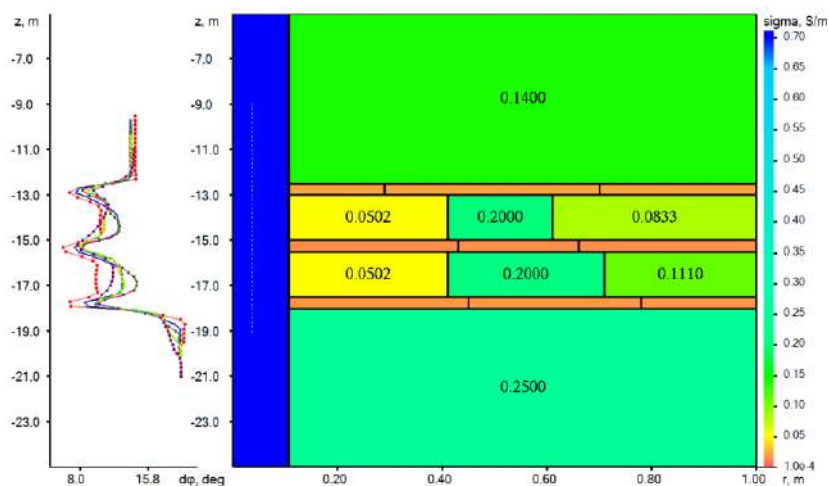


Рис. 2 – Подобранная геоэлектрическая модель среды при совместном поиске УЭП и границ блоков нефтеводонасыщенного коллектора, осложнённого карбонатными прослоями, полученная на 14-й итерации. $\Phi = 1.5E-3$.

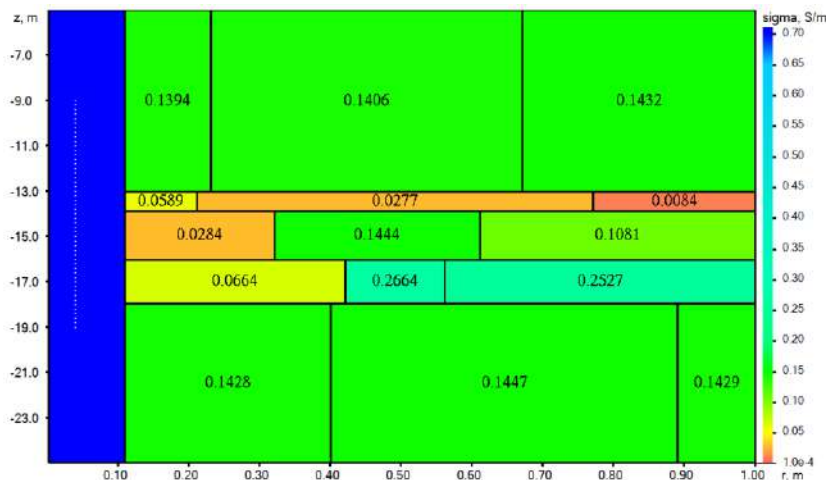


Рис. 3 – Подобранная геоэлектрическая модель при зашумленных данных, полученная на 8-й итерации. $\Phi = 3.5E-2$.

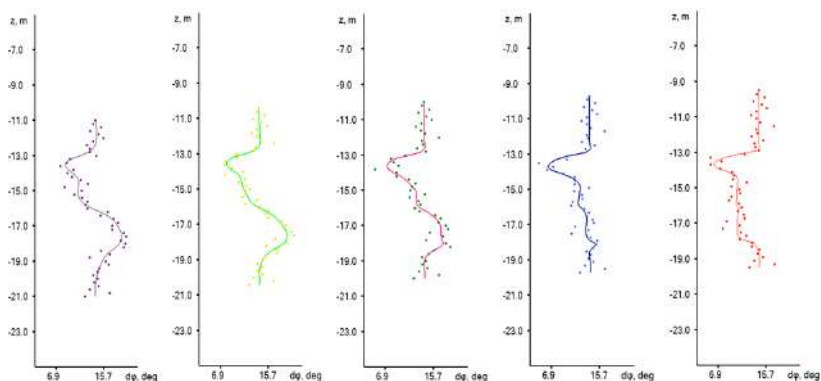


Рис. 4 – Графики зависимости разности фаз от положения зонда (сплошные линии) и зашумленные практические данные (точки).

Заключение. Был разработан программный комплекс решения прямых и обратных задач индукционного каротажа.

Верификация обратной задачи проведена путем использования синтетических аналогов полевых данных индукционного каротажа, полученных с использованием 2D-моделирования.

Проведен анализ работоспособности разработанных методов решения обратной задачи индукционного каротажа с совместным восстановлением геометрических границ и удельной электрической

проводимости. Показана хорошая сходимость к параметрам истинной модели.

Были проведены исследования качества восстанавливаемых геоэлектрических моделей в условиях зашумленных данных. Проведенные исследования показали возможность корректного восстановления геометрических границ и значений проводимости.

Литература:

1. Глинских В.Н., Ельцов И.Н., Еремин В.Н., Жмаев С.С., Каюров К.Н., Киселев В.В., Лаврухов В.Т., Мартаков С.В., Никитенко М.Н., Петров А.Н., Ревва М.Ю., Соболев А.Ю., Сухорукова К.В., Ульянов В.Н., Черяука А.Б. Технология исследования нефтегазовых скважин на основе ВИКИЗ (высокочастотное индукционное каротажное изопараметрическое зондирование). Методическое руководство. — Новосибирск: СО РАН НИЦ ОИГТМ, 2000. — 122 с.
2. Соловейчик Ю.Г., Рояк М.Э., Персова М.Г. Метод конечных элементов для решения скалярных и векторных задач. — Новосибирск: Изд-во НГТУ, 2007. — 896 с.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССА ПРОСАЧИВАНИЯ В ГЕТЕРОГЕННОЙ СРЕДЕ С КОНТРАСТНЫМИ ФИЗИЧЕСКИМИ СВОЙСТВАМИ

Д.А. Исакин

**Новосибирский государственный технический университет,
г. Новосибирск, isakin.2021@stud.nstu.ru**

Научный руководитель: Иткина Н.Б., к.т.н., доцент

Рассматривается дискретная вариационная постановка на базе разрывного метода Галеркина для задачи Дарси. Приводятся результаты численного моделирования для гетерогенной среды.

A discrete variational formulation based on the discontinuous Galerkin method for the Darcy problem is considered. The results of numerical modelling for heterogeneous medium are given.

Введение. При решении задач геофизики, связанных с добычей нефти, задач охраны окружающей среды часто требуется численное моделирование процесса просачивания в гетерогенной среде с контрастными физическими свойствами. Гетерогенная среда — это среда с включениями различной формы и размера, кавернами и когезивными зонами. Осадочные породы - характерный пример

гетерогенной среды с выраженной анизотропией физических свойств. Такая специфика области моделирования порождает достаточно жесткие требования к методу дискретизации математической модели, описывающей процесс просачивания

Один из численных методов, позволяющих учесть неконформность, как физическую, так и геометрическую — это разрывный метод Галеркина.

Постановка задачи. Рассмотрим 2D задачу просачивания в пористой среде. В форме закона Дарси она имеет вид:

$$\begin{aligned} -\nabla(K\nabla p) &= f \quad \text{в } \Omega \\ p|_{\Gamma_D} &= p_0 \quad \text{на } \Gamma_D \\ K\nabla p \cdot n &= g \quad \text{на } \Gamma_N \end{aligned} \quad (1)$$

Здесь K -симметричный положительно определенный в Ω тензор, p – давление, $\Omega \subset \mathbb{R}^2$ -ограниченная квадрируемая область, $f \in L^2(\Omega)$, $p_0 \in H^{1/2}(\Gamma_D)$ и $g \in L^2(\Gamma_N)$

Вариационная постановка. Определим конечномерные функциональные пространства, ассоциированные с разбиением $\tau_h = \{E\}$, где E – это прямоугольный конечный элемент, следующим образом:

$$\begin{aligned} V_h &= \left\{ v \in L^2(\Omega) : v|_E \in P_k(E) \quad \forall E \in \tau_h \right\} \\ \Sigma_h &= \left\{ \tau \in [L^2(\Omega)]^2 : \tau|_E \in [P_k(E)]^2 \quad \forall E \in \tau_h \right\} \end{aligned}$$

где $P_k(E)$ и $[P_k(E)]^2$ -пространство скалярных и векторных полиномиальных функций соответственно, заданных на элементе E .

Обозначим E^+ , E^- два конечных элемента с общей границей e . E^+ -правее e , E^- -левее e . n^+ , n^- - соответствующие единичные векторы внешней нормали к E^+ , E^- . Пусть ε_h – множество всех ребер. Введем операторы среднего и скачка следующим образом:

$$\begin{aligned} \{v\} &= \frac{v^+ + v^-}{2} \quad [v] = v^+ n^+ + v^- n^- \quad \forall e \in \varepsilon_h \\ \{\tau\} &= \frac{\tau^+ + \tau^-}{2} \quad [\tau] = \tau^+ n^+ + \tau^- n^- \quad \forall e \in \varepsilon_h \end{aligned} \quad (2)$$

Тогда общая вариационная формулировка в потоковой форме с учётом (2) и в соответствии с [1], запишется в следующем виде:

Найти $p_h \in L^2(\Omega): \forall v \in L^2(\Omega)$

$$\begin{aligned} \int_{\Omega} K \nabla p_h \cdot \nabla v d\Omega + \int_{\varepsilon_h} ([\hat{p}_k - p_h] \cdot \{K \nabla v\} - [v] \cdot \{K \hat{\sigma}_k\}) dS + \\ + \int_{\varepsilon_h \setminus (\Gamma_D \cup \Gamma_N)} ([\hat{p}_k - p_h] [K \nabla v] - \{v\} [K \hat{\sigma}_k]) dS = \int_{\Omega} f v d\Omega \end{aligned} \quad (3)$$

где \hat{p}_k и $\hat{\sigma}_k$ численные потоки, выбор которых влияет как на вид вариационной постановки, так и на ее свойства (устойчивость, сходимость, и т.д.). Будем использовать численные потоки, соответствующие постановки IP (Internal Penalty) [2,3]. Учитывая, что решение вариационной задачи (3) определяется в виде разложения по базисным функциям:

$$p_h = \sum_{i=0}^N q_i \psi_i \quad (4)$$

где ψ_i - базисная функция из пространства V_h ассоциированная с конечным элементом. Тогда дискретный аналог задачи (3) в форме внутреннего штрафа для функции p_h , запишется так:

$$\begin{aligned}
& \sum_{j=1}^n \left(\int_E K q_j \nabla \psi_j \cdot \nabla \psi_i dE \right) + \\
& \sum_{e \in \varepsilon_h \setminus (\Gamma_D \cup \Gamma_N)} \int_e - \left(\left[\sum_{j=1}^n q_j \psi_j \right] \cdot \{ K \nabla \psi_i \} + [\psi_i] \cdot \left\{ K \sum_{j=1}^n q_j \psi_j \right\} \right) + \\
& + \mu_e K \left[\sum_{j=1}^n q_j \psi_j \right] \cdot [\psi_i] dS + \\
& \sum_{e \in \Gamma_D} \int_e - \left(\sum_{j=1}^n q_j \psi_j \bar{n} \cdot K \nabla \psi_i + \psi_i \bar{n} \cdot K \sum_{j=1}^n q_j \nabla \psi_j \right) \\
& + \mu_e (K \bar{n} \cdot \bar{n}) \sum_{j=1}^n q_j \psi_j \psi_i dS = \int_E f \psi_i dE + \sum_{e \in \Gamma_N} \int_e g \psi_i dS + \\
& \sum_{e \in \Gamma_D} \int_e - \sum_{j=1}^n q_j \psi_j \bar{n} \cdot K \nabla \psi_i + \mu_e \psi_i p_0 (K \bar{n} \cdot \bar{n}) dS, i = 1 \dots n
\end{aligned}$$

Результаты моделирования. Представленный метод был верифицирован на модельных задачах, с контрастным тензорным коэффициентом. $K = \begin{bmatrix} 1e-12 & 0 \\ 0 & 1e-14 \end{bmatrix}$, аналитическое решение

$p = x(1-x)$ у $(1-y)$ Область моделирования $\Omega = [0,1] \times [0,1]$, шаг дискретизации равномерный $h = 0.1$. На рис.1 изображено сечение плоскостью $y = 0.2$ решения задачи. Относительная погрешность расчета в норме L^2 на 3-х вложенных сетках составляет соответственно: $\Delta_h = 8.5e-4$, $\Delta_{h/2} = 2.1e-4$, $\Delta_{h/4} = 5.4e-5$.

Заключение. По результатам вычислительных экспериментов можно сделать вывод, что представленная схема, довольно точно описывает процесс просачивания в гетерогенной среде, даже на грубой сетке, с контрастными физическими свойствами.

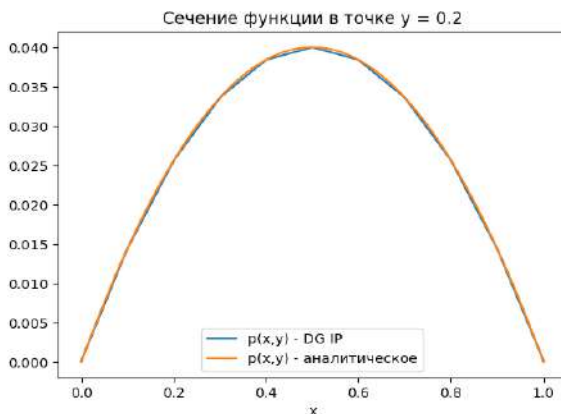


Рис.1 – сравнение поля давления истинного и рассчитанного

Литература:

1. Arnold D.N., Brezzi F, Cocburn B., Marini D. Unified analysis of discontinuous Galerkin methods for elliptic problems //SIAM J. Numer. Anal. 2002. V.39. №5. P.1749-1779.
2. Girault V., Riviere B., Wheeler M.F. A discontinuous Galerkin method with non-overlapping domain decomposition for the Stokes and Navier-Stokes problems// Mathematics of Computation. 2005. No. 74. P. 53-84.
3. Burman E., Hansbo P. A unified stabilized method for Stokes and Darcy's equations // J. of Computational and Applied Mathematics. 2007. No. 198. P. 35-51.

АНАЛИЗ ЭФФЕКТИВНОСТИ ПОЛИМЕРНОГО ЗАВОДНЕНИЯ С ИСПОЛЬЗОВАНИЕМ ПРОЦЕДУРЫ ОПТИМИЗАЦИИ

Е.А. Кравец, М.Г. Персова

**Новосибирский государственный технический университет,
г. Новосибирск, katherinkravets@mail.ru**

Научный руководитель: Персова М.Г., д.т.н., профессор

Разработан программный модуль, который позволяет оптимизировать режимы разработки нефтяных месторождений. Для модели коллектора с неоднородной проницаемостью проведены исследования эффективности полимерного заводнения.

A software module has been developed that allows for adjustment of oil field development modes. For a reservoir model with non-uniform permeability, studies of the efficiency of polymer flooding were conducted.

Добыча нефти является одной из важных отраслей промышленности, которая обеспечивает энергией многие страны мира. Однако в процессе добычи нефти возникает проблема выбора применения различных технологий и методов с целью повышения экономической эффективности разработки. Для этого применяются методы оптимизации, основанные на гидродинамическом моделировании для цифровых моделей месторождений [1-2]. Предварительная цифровая модель (перед началом добычи) строится на основании результатов геологоразведки и данных каротажа скважин, а затем эта модель постепенно адаптируется (уточняется) на основе поступающих добычи путем решения обратных задач и гидродинамического моделирования [3]. Обычно на начальном этапе добычи нефть за счет естественного внутрипластового давления через скважины фонтанирует на поверхность фактически самостоятельно. Поэтому сначала при добыче из пласта смеси используются первичные методы, когда затраты на добычу нефти являются минимальными. Со временем пластовое давление снижается, нефть перестает фонтанировать из скважины, и первичный метод добычи теряет свою эффективность. Так как много нефти все еще остается в пласте, необходимо использовать вторичные и третичные методы. На этом этапе используются, например, технологии полимерного заводнения, которые позволяют изменять свойства флюидов и пластовых пород [4]. Концентрация закачиваемых агентов играет важную роль в этом процессе, так как она влияет на вязкость воды, используемой для закачки в скважину [4].

Функционал, который минимизируется в процессе оптимизации режимов работы скважин, имеет следующий вид [2]:

$$\begin{aligned} \Phi(\theta) = & \sum_{w=1}^{N_w} \left(\int_t a_1 V_w^{water,prod}(\theta) dt \right)^2 + \sum_{w=1}^{N_w} \left(\int_t a_2 M_w^{poly}(\theta) dt \right)^2 + \\ & + \sum_{w=1}^{N_w} \left(\int_t a_3 V_w^{water,inj}(\theta) dt \right)^2 + \sum_{w=1}^{N_w} \left(\int_t a_4 |V_w^{oil,*}(\theta) - V_w^{oil,*}(\theta)| dt \right)^2 + \\ & + \sum_{r=1}^{N_r} \gamma_r (\theta_r - \bar{\theta}_r)^2 \rightarrow \min_{\theta}, \end{aligned} \quad (1)$$

где N_w — количество скважин (нагнетательные или добывающие), $V_w^{water,prod}$ — объем воды в добывающей скважине, M_w^{poly} — масса закаченного полимера в нагнетательную скважину, $V_w^{water,inj}$ — объем закаченной воды в нагнетательную скважину, $V_w^{oil,*}$ — количество нефти, которую требуется извлечь, a_k — подобранные коэффициенты пользователем, в зависимости от стоимости нефти и затрат на полимер.

Для удобства распишем данную формулу следующим образом:

$$F_{1,w}(\theta) = \int_t (a_1 V_w^{water,prod}(\theta) dt)^2, F_{2,w}(\theta) = \int_t (a_2 M_w^{poly}(\theta) dt)^2,$$

$$F_{3,w}(\theta) = \int_t (a_3 V_w^{water,inj}(\theta) dt)^2, F_{4,w}(\theta) = \int_t (a_4 |V_w^{oil} - V_w^{oil,*}| dt)^2.$$

Выполним с помощью разложения в ряд Тейлора линейаризацию $F_{k,w}(\theta)$ в окрестности параметров θ_i . Отбросим все слагаемые после первой производной и получим следующее:

$$F_{k,w}(\theta) \approx F_{k,w}(\bar{\theta}) + \sum_{r=1}^{N_r} \left. \frac{\partial F_{k,w}}{\partial \theta_r} \right|_{\theta = \bar{\theta}} \Delta \theta_r, k = 1 \dots 4, \quad (2)$$

где N_r – размер вектора параметров.

Подставим соотношение (2) в функционал (1) и приравняем производные функционала по параметрам к нулю. Откуда получим СЛАУ вида $(A + \gamma) \Delta \theta = b$, где матрица A и вектор b считаются следующим образом:

$$A_{ij} = \sum_{k=1}^4 \sum_{w=1}^{N_w} \frac{\partial F_{k,w}}{\partial \theta_i} \frac{\partial F_{k,w}}{\partial \theta_j}, i, j = 1 \dots N_r, \quad (3)$$

$$b_i = - \sum_{k=1}^4 \sum_{w=1}^{N_w} F_{k,w}(\bar{\theta}) \frac{\partial F_{k,w}}{\partial \theta_i}. \quad (4)$$

Производные в (1) и (2) будем искать следующим образом:

$$\frac{\partial F_{k,w}}{\partial \theta_i} = \frac{\delta F_{k,w}(\theta_i + \Delta \theta_i) - \delta F_{k,w}(\theta_i)}{\Delta \theta_i}, \quad (5)$$

где $\Delta \theta_i = 0.2 * \theta_i$ – приращение к параметру θ_i , взятое как 20% от его значения.

Проведены исследования на двух моделях – с низкопроницаемой подобластью на нефтяном месторождении и без нее (Рисунок 1). При рассмотрении каждой модели нефтяного пласта были установлены весовые коэффициенты a_k в зависимости от того, какой план необходимо было найти – с умеренной экономией полимера или со значительной экономией. Для каждого плана эксплуатации скважин были рассчитаны прибыль и затраты на полимер. Эти расчеты позволили оценить экономическую эффективность каждого плана и выбрать тот, который обеспечивает оптимальное сочетание добытой нефти и затраченного полимера.

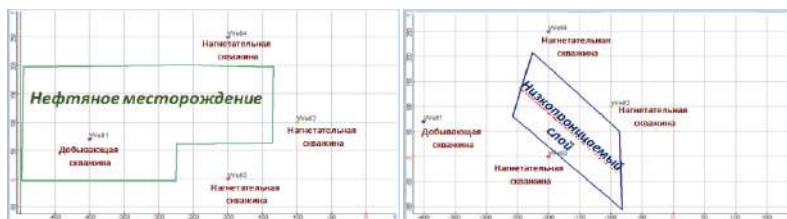


Рисунок 1 – Подобласти месторождений

Рассмотрим исследование, где производилась оптимизация режима работы всех скважин для модели с низкопроницаемой подобластью. Для расчетов были установлены следующие показатели: цена полимера за 1 килограмм составляла 406.4 рублей, цена за 1 тонну нефти составляла 19183 рубля [4]. Результаты представлены в таблице 1 и таблице 2. При слабой экономии полимера использовались весовые коэффициенты $a_1 = 0$, $a_2 = 2$, $a_3 = 0.005$, $a_4 = 0.1$, а при сильной экономии полимера – $a_1 = 0$, $a_2 = 5$, $a_3 = 0.4$, $a_4 = 0.1$.

Из представленных результатов видно, что большую эффективность обеспечивает план с сильной экономией полимера, поскольку он позволяет получить практический такой же прирост нефти, что и план, полученный при слабой экономии полимера, но при меньших затратах на полимер.

Таблица 1 – Показатели нефти и полимера при стартовом и оптимальном плане

	Добытая нефть, т.	Полимер, кг.
До синтеза без полимера	7717.92	0.00
До синтеза с полимером	9317.53	70.08
Слабая экономия полимера	9495.84	69.05
Сильная экономия полимера	9486.00	55.65

**Таблица 2 – Расчет затрат и прибыли при стартовом
и оптимальном плане**

	Доп. доб. нефть	При- быль за доп. доб. нефть, млн. руб	Прибыль за нефть без доп. доб., млн. руб	Суммар- ная прибыль, млн. руб	Цена за поли- мер, млн. руб.	Итог. При- быль, млн. руб.
До синтеза без полимера	0.000	0.000	148.053	148.053	0.000	148.053
До синтеза с полимером	1599.6 12	30.685	148.053	178.738	28.481	150.258
Слабая экономия полимера	1777.9 21	34.106	148.053	182.159	28.060	154.098
Сильная экономия полимера	1768.0 80	33.917	148.053	181.970	22.614	159.356

Литература:

1. Xue X., Chen G., Zhang K., et al. A divide-and-conquer optimization paradigm for waterflooding production optimization // Journal of Petroleum Science and Engineering. Elsevier, 2022.
2. Персова М.Г., Соловейчик Ю.Г., Патрушев И.И., Насыбуллин А.В., Алтынбекова Г.Ж., Леонович Д.А. Оптимизация разработки нефтяных месторождений на основе 3D-модели пласта, полученной в результате автоадаптации // Известия Саратовского университета. Новая серия. Серия: Математика. Механика. Информатика. – 2023. – Т. 23, № 4. – С. 544–558. – doi 10.18500/1816-9791-2023-23-4-544-558.
3. Persova M.G., Soloveichik Y.G., Vagin D.V., Grif A.M., Kiselev D.S., Patrushev I.I., Nasybullin A.V., Ganiev B.G. The design of high-viscosity oil reservoir model based on the inverse problem solution // Journal of Petroleum Science and Engineering. – 2021. – Vol. 199. – Art. 108245. <https://doi.org/10.1016/j.petrol.2020.108245>.
4. Насыбуллин А.В., Персова М.Г., Орехов Е.В., Шайдуллин Л.К., Соловейчик Ю.Г., Патрушев И.И. Моделирование ПАВ-полимерного заводнения на участке Бурейкинского месторождения // Нефтяное хозяйство. – 2022. – № 7. – С. 38–42. – doi 10.24887/0028-2448-2022-7-38-42.

МОДЕЛИРОВАНИЕ МИКРОСОСТОЯНИЙ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ ОПИСАНИЯ МАКРОСКОПИЧЕСКИХ ФИЗИКО-ХИМИЧЕСКИХ ПРОЦЕССОВ

Н.К. Лысенко

Томский политехнический университет,

г. Томск,

Научный руководитель: В. М. Саклаков

Многие физико-химические процессы моделируются на уровне макроскопических систем, что может накладывать ограничение на точность получаемых результатов. В настоящей работе проводится эксперимент по образованию фторида лития в результате реакции ионного обмена между гидроксидом лития и фторидом аммония. Для анализа результатов применяется инструментарий физической химии и квантовой механики. Эксперимент показал частоту, с которой на микроуровне возникают эффекты, не характерные для макроуровня.

Many physical and chemical processes are modeled at the level of macroscopic systems, which can impose a limitation on the accuracy of the results. In this paper, an experiment is conducted on the formation of lithium fluoride as a result of ion exchange reaction between lithium hydroxide and ammonium fluoride. The tools of physical chemistry and quantum mechanics are used. The experiment showed frequency with which effects occur at the micro level that are not characteristic of the macro level.

Известно, что объектом исследования физической химии являются дисперсные макроскопические системы. В связи со статистическим характером её положений, интегральной природой функций состояния, необходимым допущением является отклонение от равномерного максвелловского распределения, а кроме того нарушение на уровне микросостояний условий протекания самопроизвольных процессов (убыль изобарно-изотермического потенциала G , H – энтальпия процесса, S – энтропия процесса, U – свободная внутренняя энергия, P – давление, V – объём) [1-2, 4].

$$G \equiv U + PV - TS$$

$$\Delta G = \Delta U + P\Delta V + P\Delta V - \Delta TS - T\Delta S = \Delta U + P\Delta V - T\Delta S$$

$$H \equiv U + PV$$

$$\Delta H = \Delta U + P\Delta V + P\Delta V = \Delta U + P\Delta V$$

$$\Delta G = \Delta H - T\Delta S < 0$$

Кристаллизация как фазовый переход первого рода, с одной стороны являясь самопроизвольным неравновесным процессом, сопровождается приращением энтропии и снижением изобарно-изотермического

потенциала системы, выравниванием его величины и величины его дифференциала между отдельными частями системы (фазой маточного раствора и образующихся центров кристаллизации), а с другой стороны требует для зарождения самих центров кристаллизации локального увеличения плотности вследствие флуктуации концентраций при диффузионном переносе вещества по объёму системы, что можно трактовать как местное снижение энтропии, так как повышается упорядоченность упаковки частиц дисперсной фазы [1-3].

При этом известно, что в процессе формирования поверхности границы раздела фаз, избыточную энергию поверхностного натяжения компенсируется убылью того избытка энергии Гиббса, который был обусловлен пересыщенным состоянием раствора ($(\frac{\delta G}{\delta n_i})_{P,T} = \mu$ - химический потенциал, $\sigma = (\frac{\delta G}{\delta S})_{P,T}$ - поверхностное натяжение, S - площадь поверхности границы раздела фаз) [3-4]

$$\Delta G_{\text{система}} = \Delta G_V + \Delta G_S = (\frac{\delta G}{\delta n_i})_{P,T} n_i + \sigma S = n(\mu_{\text{ж}} - \mu_{\text{тв}}) + \sigma S < 0$$

$$\sigma S > 0, \quad \mu_{\text{ж}} > \mu_{\text{тв}}, \quad n(\mu_{\text{ж}} - \mu_{\text{тв}}) < 0$$

Избыточная поверхностная энергия отдельных образующихся кластеров, центров кристаллизации, будет снижаться путем уменьшения удельной площади поверхности ($S_{\text{уд}}$), то есть по мере укрупнения до размеров коллоидных частиц ($S_{\text{пов}} = 4\pi r^2$ - площадь поверхности сферической частицы с объёмом $V = \frac{4}{3}\pi r^3$ и радиусом r) [3].

$$S_{\text{уд}} = \frac{S_{\text{пов}}}{V} = \frac{4\pi r^2}{\frac{4}{3}\pi r^3} = \frac{3}{r} < 0 \Leftrightarrow r > 0$$

Таким образом, статистически система в результате образования суспензии приходит к метастабильному состоянию динамического равновесия, которым характеризуется насыщенный раствор. Однако путь протекания этого процесса может включать процессы, как приближающие, так и отдаляющие систему от равновесия [1-2].

Целью данного исследования является применение объединенного математического аппарата физической химии и квантовой механики для

более детального описания макроскопических физико-химических процессов. Объектом исследования является процесс образования суспензии фторида лития в результате реакции ионного обмена между гидроксидом лития и фторидом аммония.

Моделирование процесса образования суспензии фторида лития было выполнено в несколько этапов. На основании ограничений, накладываемых на самопроизвольные процессы в термодинамических системах, а также эмпирических данных по кинетике реакции, были подобраны случайные совокупности взаимодействий между частицами на уровне отдельных микросостояний, удовлетворяющие макроскопическим ограничениям. Затем расчеты в рамках классической механики, связанные с взаимодействиями отдельных частиц, были уточнены с помощью положений квантовой механики.

Многоуровневая структура модели отвечает постепенно возрастающей, по мере приближения к описанию микросостояний, сложности вычислительных задач. Она разделена на отдельные модули: математический аппарат физической химии с функциями состояния и их дифференциалов при постоянстве ряда параметров системы, модуль с уравнениями из электростатики и электродинамики, выражениями химической кинетики, коллоидной химии, поверхностных явлений, и наконец квантовой химии. В отдельный модуль вынесены функции, составляющие случайные совокупности процессов на уровне отдельных микросостояний, ограниченные по суммарному вкладу в изменение функций состояния системы.

Скорость химической реакции и её порядок определяли спектрофотометрически, поскольку в определенном диапазоне концентраций обратная зависимость оптической плотности раствора от концентрации соединения, образующего коллоидные частицы, обладает линейным характером. Для определения этой градуировочной зависимости, были приготовлены коллоидные растворы фторида лития, причем оптическую плотность измеряли сразу после образования суспензии. Для проведения кинетических расчетов, необходимо было определить скорость и её изменение от времени химической реакции – константа скорости думолекулярной реакции образования фторида лития была рассчитана по изменению оптической плотности, то есть от концентрации коллоидных частиц. Второй порядок реакции был установлен по характерному соотношению времени превращения половины ко времени превращения четверти количества реагентов.

Суспензии фторида лития характеризуются устойчивым седиментационно-диффузионным равновесием, что указывает на прерывание роста коллоидных частиц при достижении определенного

размера, препятствующее образованию осадка. Предельный размер коллоидных частиц был определен в соответствии с гипсометрическим законом Лапласа-Перрена по приблизительной величине плотности, объема частиц и их числу, установленному спектрофотометрически.

При моделировании образования коллоидных частиц так же принимали во внимание равновесные концентрации реагентов и продуктов реакции, определенные в соответствии с константой равновесия взаимодействия гидроксида лития с фторидом аммония.

Известно, что при малой скорости осаждения над ней преобладает скорость процесса ориентационных кулоновских взаимодействий между диполями молекул, образующих коллоидные частицы, что приводит к получению кристаллического осадка. Это обстоятельство позволило применить для описания квантово-механических взаимодействий электронных оболочек ионов лития и фтора энергию молекулярных орбиталей на основании спектров поглощения кристаллического фторида лития.

Многочисленное повторение моделирования позволило определить частоту, с которой на уровне микросостояний возникают не характерные для макроскопического состояния изменения функций состояния системы. Кроме того, была определена величина их среднеквадратичных отклонений.

Полученные с помощью моделирования данные показывают увеличение точности описания физико-химического процесса при росте размера выборки микропроцессов. Данный фактор накладывает технологические ограничения, связанные с предельным доступным объемом оперативной памяти вычислительной машины.

В результате исследования было установлено: при наличии достаточных количественных характеристик процессов, протекающих на микроуровне, возможно моделировать микросостояния. Это позволяет уточнять представление о конкретных физико-химических процессах при определенных условиях. Данный метод может быть применен в ряде практических областей, например, при разработке структуры поверхности сорбентов для хроматографических систем.

Литература

1. Герасимов Я. И., Древинг В. П., Еремин Е. Н., Киселев А. В., Лебедев В. П., Панченков Г. М., Шлыгин А. И. Курс физической химии. В 2 т. Т. 1. – М.: Издательство «Химия», 1964 – 624 с.
2. Глазов В. М. Основы физической химии. – М.: Высшая школа, 1981 – 456 с.

3. Фролов Ю.Г. Курс коллоидной химии. Поверхностные явления и дисперсные системы. Учебник для вузов. - 2-е изд. – М.: Химия, 1988 – 464 с.

4. Зломанов В. П., Казин П. Е., Яценко А. В. Вещество и его превращения: основные понятия // Конденсированные среды и межфазные границы. – 24 (2) – 2022. – С. 211-219.

ЧИСЛЕННАЯ МОДЕЛЬ ИНДУКЦИОННОГО НАГРЕВА ПОД ЗАКАЛКУ С УЧЕТОМ ПРОЦЕССА АУСТЕНИЗАЦИИ В СТАЛИ

Ф.В. Чмиленко¹, И.И. Растворова², В.Е. Парменов¹

**¹Санкт-Петербургский государственный электротехнический
университет «ЛЭТИ» им. В.И.Ульянова (Ленина),**

**²Санкт-Петербургский горный университет
императрицы Екатерины II,**

г. Санкт-Петербург, parmenov.slava@yandex.ru

Научный руководитель: Чмиленко Ф.В., д.т.н., профессор

Работа посвящена моделированию индукционной поверхностной закалки. Задачей настоящей работы является разработка алгоритма и численной модели на его основе, которая определяет требуемую удельную поверхностную мощность и время нагрева для деталей цилиндрической формы при заданных частоте, диаметре детали и глубине закалки с учетом зависимостей температур фазовых превращений на поверхности и на глубине закалки от скорости нагрева в этих зонах детали.

The work is devoted to induction surface hardening modeling. The task of this work is to develop an algorithm and a numerical model based on it, which determines the required specific surface power and heating time for cylindrical-shaped parts at a given frequency, part diameter and hardening depth, taking into account the phase transformations temperature dependence on the surface and at the hardening depth on the heating rate in these areas of the part.

Известно, что при индукционной поверхностной закалке на формирование закаленного слоя влияет большое количество разнообразных факторов. Основные из них это: температура и скорость нагрева, химический состав и исходная структура, теплофизические свойства стали, условия внешнего и внутреннего теплоотвода. На этапе нагрева в значительной степени качество термической обработки определяется точностью фиксации моментов аустенизации и гомогенизации структуры металла [1, 2]. В свою очередь температуры аустенизации и гомогенизации зависят от скорости нагрева.

Информация о зависимости от скорости нагрева температуры критических точек $Ac1$ и $Ac3$, а также температуры гомогенизации T_H для стали 45 была взята из термокинетической диаграммы аустенизации [3] (см. рис. 1). Скорость нагрева вычислялась начиная с минимальной температуры начала аустенизации $T_{Ac1\text{ мин}}$.

Корректная работа с термокинетической диаграммой возможна только вдоль линии постоянной скорости нагрева. Данная диаграмма не приспособлена для учета времени выдержки.

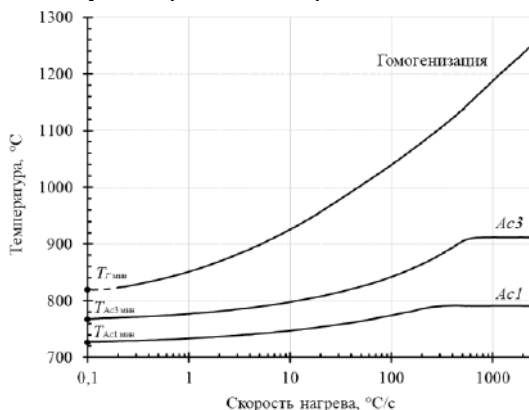


Рис.1 – Зависимости температур начала фазовых превращений ($Ac1$), аустенизации ($Ac3$) и гомогенизации от скорости нагрева для стали 45.

Для поверхности время выдержки определяется задержкой, когда заготовка уже не греется, а интенсивное охлаждение еще не началось. Для глубины закалки дополнительно характерна еще и естественная выдержка, вызванная ограниченной скоростью теплоотвода.

Для корректного использования термокинетических диаграмм аустенизации, авторы применили понятие эквивалентной скорости нагрева. Для ее расчета требуется максимальный перепад температур поделить на время, которое включает фазу нагрева и фазу охлаждения до минимальной температуры гомогенизации $T_{H\text{ мин}}$ для поверхности или до минимальной температуры конца аустенизации $T_{Ac3\text{ мин}}$ для глубины закалки.

В работе [4] показано, что электротепловые модели, основанные на расчете по первой гармонике, дают большую погрешность при моделировании нагрева под закалку, поэтому электротепловой расчет проводился только во временной области, по численной модели, описанной в [5].

Блок-схема итерационного поиска поверхностной удельной мощности $P_{y\phi}$ и времени нагрева t_n под закалку приведена на рис. 2.

Алгоритм итерационного поиска состоит из следующих этапов:

1. Ввод основных данных для расчета: марка стали, диаметр детали, частота тока, глубина закалки, коэффициент теплоотдачи α и время задержки после окончания нагрева и до начала охлаждения. Ввод предполагаемых температуры нагрева поверхности T_n и $P_{y\phi}$. Чем ближе предполагаемые значения к искомым, тем меньше итераций потребуется алгоритму.

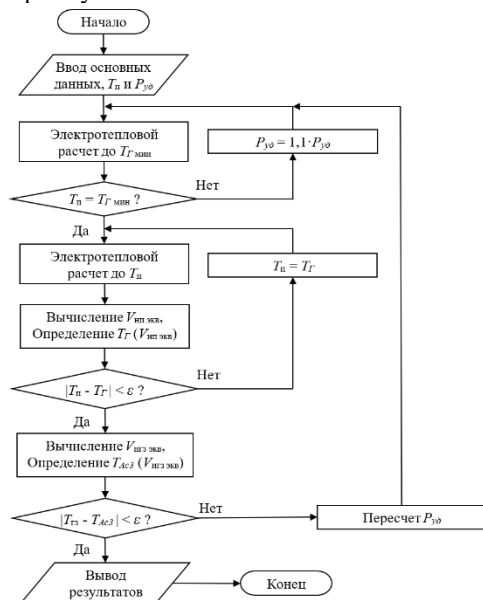


Рис.2 – Блок-схема поиска удельной мощности и времени нагрева

2. Выполняется электротепловой расчет с попыткой нагрева поверхности до $T_{Г\text{ мин}}$.

3. Если $P_{y\phi}$ оказалась недостаточной, для достижения $T_{Г\text{ мин}}$, она увеличивается на 10% и электротепловой расчет п.2 выполняется заново. При достижении $T_{Г\text{ мин}}$ на поверхности, алгоритм идет дальше.

4. Выполняется электротепловой расчет нагрева с температуры, полученной в п.2 до T_n . Здесь же моделируется задержка перед охлаждением и стадия интенсивного охлаждения.

5. Вычисляется эквивалентная скорость нагрева поверхности $V_{\text{нп экв}}$. По термокинетической диаграмме определяется температура гомогенизации T_G , соответствующая $V_{\text{нп экв}}$.

6. Проводится проверка равенства T_n и определенной T_r . Если ошибка равенства больше заданной, то T_n корректируется и электротепловой расчет п.4 выполняется заново. При успешной проверке равенства, алгоритм идет дальше.

7. Вычисляется эквивалентная скорость нагрева на глубине закалки $V_{нгз экв}$. По термокинетической диаграмме определяется температура аустенизации T_{Ac3} , соответствующая $V_{нгз экв}$.

8. Проводится проверка равенства $T_{гз макс}$ и определенной T_{Ac3} . Если ошибка равенства больше заданной, то $P_{уд}$ корректируется методом секущих и электротепловой расчет п.2 выполняется заново. При успешной проверке равенства, работа алгоритма завершается.

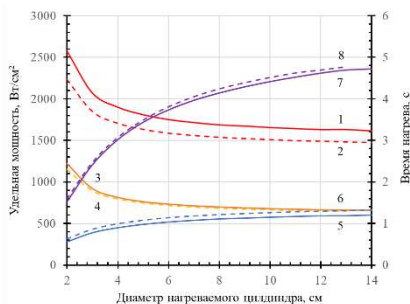
9. Рассчитываются и сохраняются основные и вспомогательные результаты: $P_{уд}$, t_n ; T_r и $V_{нп экв}$ для поверхности детали; T_{Ac3} и $V_{нгз экв}$ для глубины закалки.

На рис. 3 показаны результаты расчетов $P_{уд}$ и t_n , полученные для глубин закалки 0,2 и 0,3 см и временах выдержки 0,1 и 0,4 с.

На рис. 4 показаны результаты расчетов T_r и $V_{нп экв}$ для поверхности детали, T_{Ac3} и $V_{нгз экв}$ для глубины закалки, полученные для глубин закалки 0,2 и 0,3 см при времени выдержки 0,1 с.

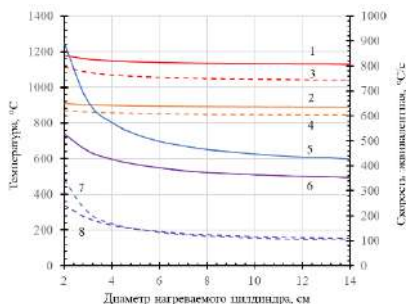
Все результаты получены для цилиндрических деталей из стали 45 при частоте тока 10 00 Гц и $\alpha = 40\,000 \text{ Вт/м}^2 \cdot \text{К}$.

Разработанная модель позволяет находить оптимальные параметры термобработки (время нагрева и удельную поверхностную мощность) с учетом влияния скорости нагрева на процесс аустенизации в стали, также предоставляя дополнительную информацию о технологическом процессе. Приведенный алгоритм можно легко приспособить для анализа индукционного нагрева других углеродистых сталей, для которых доступна термокинетическая диаграмма аустенизации. Дальнейшее повышение точности разработанной модели может быть достигнуто на основе валидации и внесения для каждой используемой стали своих поправочных коэффициентов.



- | | |
|-----------------------------|--------------------------|
| 1 – 0.2 см, 0.1 с, $P_{уд}$ | 5 – 0.2 см, 0.1 с, t_n |
| 2 – 0.2 см, 0.4 с, $P_{уд}$ | 6 – 0.2 см, 0.4 с, t_n |
| 3 – 0.3 см, 0.1 с, $P_{уд}$ | 7 – 0.3 см, 0.1 с, t_n |
| 4 – 0.3 см, 0.4 с, $P_{уд}$ | 8 – 0.3 см, 0.4 с, t_n |

Рис.3 – Результаты расчетов зависимостей Руд и тн от диаметра нагреваемого цилиндра на частоте тока 10 000 Гц для стали 45



- | | |
|-------------------------|---------------------------|
| 1 – 0.2 см, $T_{гн}$ | 5 – 0.2 см, $V_{нп экв}$ |
| 2 – 0.2 см, $T_{Ас3гз}$ | 6 – 0.2 см, $V_{нгз экв}$ |
| 3 – 0.3 см, $T_{гн}$ | 7 – 0.3 см, $V_{нп экв}$ |
| 4 – 0.3 см, $T_{Ас3гз}$ | 8 – 0.3 см, $V_{нгз экв}$ |

Рис.4 – Зависимости ТГ и $V_{нп экв}$ для поверхности, $T_{Ас3}$ и $V_{нгз экв}$ для глубины закалки от диаметра нагреваемого цилиндра на частоте тока 10 000 Гц для стали 45

Литература:

1. Зимин Н.В. О влиянии температуры, скорости нагрева и исходного состояния структуры углеродистых сталей на процессы образования в них аустенита, Металлообработка. – 2006. – № 1 (31).
2. Безменов Ф.В., Зимин Н.В. Контроль процесса аустенизации при индукционном нагреве под поверхностную закалку цилиндрических объектов, Металлообработка. – 2010. – № 6 (60).
3. Orlingh J., Wiest A.R.P. Atlas zur Wärmebehandlung der Stähle. Band 3. Dusseldorf: Verlag Stahleisen M.B.H., 1973. – P. 264.
4. Чмиленко Ф.В., Шанин Ю.В., Хоршев А.А., Бондарь А.С., Чжан Ци, Бондаренко Д.Н. Моделирование технологических параметров индукционной поверхностной закалки на основе обобщенных экспериментальных исследований / // Известия СПбГЭТУ «ЛЭТИ» 2022. – Т.15, № 2, – С. 67-78.
5. Демидович В.В., Чмиленко Ф.В. Численные методы в теории индукционного нагрева: монография. СПб.: ООО «Технолит», 2008. 220 с.

ПРИМЕНЕНИЕ МЕТОДА БАЛАНСИРОВКИ ПОТОКОВ ПРИ РЕШЕНИИ ЗАДАЧИ ФИЛЬТРАЦИИ МЕТОДОМ КОНЕЧНЫХ ЭЛЕМЕНТОВ НА НЕСОГЛАСОВАННЫХ СЕТКАХ С УЧЁТОМ ГРАВИТАЦИИ

М.Д. Романов

Новосибирский государственный технический университет,
г. Новосибирск, m.romanov.2019@stud.nstu.ru

Научный руководитель: Соловейчик Ю.Г., д.т.н., профессор

В работе приведен новый способ применения метода балансировки потоков при решении задачи фильтрации с помощью МКЭ на несогласованных сетках. Проведены вычислительные эксперименты без метода балансировки потоков, с методом балансировки потоков и с предложенной модификацией метода балансировки потоков на несогласованных сетках. Результаты вычислительных экспериментов показали, что для задач фильтрации с учётом гравитации предложенная модификация метода оказалась более эффективной.

The paper presents a new way to apply the flow balancing method in solving the filtering problem using FEM on non-conforming meshes. Computational experiments were carried out without the flow balancing method, with the flow balancing method and with the proposed modification of the flow balancing method on non-conforming meshes. The results of computational experiments have shown that the proposed modification of the method proved to be more effective for filtration problems taking into account gravity.

Математическая модель. Для несжимаемых изотермических и неперемешиваемых фаз плотность является константой. Тогда, применив к закону Дарси закон сохранения масс и некоторые алгебраические преобразования, получим следующую краевую задачу:

$$-\operatorname{div} \left(\mathbf{K} \sum_{m=1}^M \frac{\kappa^m}{\eta^m} \operatorname{grad} (P + P_c^m + \rho^m g z) \right) = 0, \quad (1)$$

$$P|_{\Gamma^1} = P^\Gamma, \quad (2)$$

$$\mathbf{K} \sum_{m=1}^M \frac{\kappa^m}{\eta^m} \left(\operatorname{grad} (P + P_c^m + \rho^m g z) \right) \Big|_{\Gamma^2} \cdot \vec{n} = F^\Gamma, \quad (3)$$

где $\mathbf{K}(x, y, z, t)$ – тензор структурной проницаемости матрицы–породы, $\kappa^m(S^m)$ – индивидуальный множитель к структурной проницаемости матрицы–породы, η^m – вязкость m -ой

фазы, $P(x, y, z, t)$ – давление, P_c^m – капиллярное давление для m -ой фазы, $\rho^m(P)$ – плотность m -ой фазы, g – ускорение свободного падения, z – координата на вертикальной оси, $S^m(x, y, z, t)$ – насыщенность m -ой фазы, Γ^1 – границы расчетной области, на которых задано давление P^Γ , Γ^2 – непроницаемые границы и зоны перфорации скважин, на которых F^Γ определяет удельный поток смеси.

Краевая задача (1)-(3) решается с помощью МКЭ.

Расчёт и постобработка численных потоков фаз. Так как МКЭ не гарантирует неразрывность производных численного решения, то для получения потоков фаз через грани вычисляется взвешенное значение потоков фаз через одну грань со стороны двух смежных элементов. Причём для вычисления потока берётся коэффициент κ^m/η^m с элемента, с которого поток вытекает. Для внешних граней вычисляется обычный поток.

МКЭ также не гарантирует выполнения закона сохранения масс. Для повышения точности выполнения закона сохранения масс применяется *метод балансировки потоков* (МБП). Суть метода состоит в вычислении корректирующих добавок δV_{Γ_i} к потокам смесей через грань Γ_i с помощью минимизации методом наименьших квадратов функционала

$$\sum_{e=1}^{N^e} \beta_e \left(\sum_{i \in I_{\Omega_e}} \left(\tilde{\gamma}_{\Omega_e, \Gamma_i} \cdot |V_{\Gamma_i}| + \gamma_{\Omega_e, \Gamma_i} \cdot \delta V_{\Gamma_i} \right) \right)^2 + \sum_{i=1}^{N^f} \alpha_i \left(\delta V_{\Gamma_i} \right)^2 \rightarrow \min_{\delta V_{\Gamma_i}}, \quad (4)$$

где N^e и N^f – число конечных элементов и граней, I_{Ω_e} – множество номеров граней элемента Ω_e , β_e и α_e – параметры регуляризации, V_{Γ_i} – осредненный поток смеси через грань Γ_i , величины $\tilde{\gamma}_{\Omega_e, \Gamma_i}$ и $\gamma_{\Omega_e, \Gamma_i}$ определяются следующим образом:

$$\tilde{\gamma}_{\Omega_e, \Gamma_i} = \begin{cases} -1, & \text{если } V'_{\Gamma_i} \text{ втекает в } \Omega_e, \\ 1, & \text{иначе.} \end{cases}, \quad (5)$$

$$\gamma_{\Omega_p, \Gamma_i} = \vec{n}_{\Gamma_i} \cdot \vec{n}_{\Omega_p, \Gamma_i}, \quad (6)$$

где \vec{n}_{Γ_i} – вектор нормали к грани Γ_i , независящей от выбора элемента, $\vec{n}_{\Omega_e, \Gamma_i}$ – вектор внешней нормали к грани Γ_i элемента Ω_e .

С помощью полученных корректирующих добавок пересчитываются потоки фаз так, чтобы их доля осталась неизменной.

МБП на несогласованных сетках. В несогласованных сетках [1] грани одного элемента могут примыкать несколько других элементов. Грани, к которым примыкает несколько элементов, будем называть «большими», а грани, на которые делятся «большие», будем называть «малыми». Остальные грани будем называть «обычными».

В работе [2] для применения МБП на несогласованных сетках предлагается учитывать только «обычные» и «большие» грани. В таком случае корректирующие добавки для «малых» граней масштабируются на основании отношения площадей «большой» и «малой» граней. В этом варианте МБП потоки на всех «малых» гранях, примыкающих к одной «большой» имеют одно направление, что в общем случае неверно. При учете гравитации потоки на «малых» гранях могут быть направлены в разные стороны из-за разных свойств одной фазы в разных ячейках.

Чтобы допустить разнонаправленные потоки на «малых» гранях предлагается учитывать только «обычные» и «малые» грани. В таком случае корректирующая добавка на «большой» грани будет суммой корректирующих добавок на примыкающих «малых» гранях.

Вычислительные эксперименты. Для определения эффективности предложенного метода были проведены два вычислительных эксперимента, в которых сравнивались три решателя: решатель (P1), основанный только на МКЭ, решатель (P2), основанный на МБП с «большими» гранями, и решатель (P3), основанный на МБП с «малыми» гранями. Вычислительные эксперименты проводились для цифровых моделей нефтяных месторождений (M1 и M2) на последовательности вложенных сеток. Вид сверху базовых конечно-элементных сеток представлен на рисунке 1.

На рисунке 2 представлены графики суммарного отбора нефти для M1 и M2, рассчитанные с помощью трёх решателей. Как видно из рисунка 2, графики, полученные с помощью P1, сходятся к графикам, полученным с помощью P2 и P3, которые, в свою очередь, оказались очень близкими на вложенных сетках. При этом графики, полученные с помощью P3, меняются слабее, чем графики P2, и графики P2 сходятся к графикам P3. Таким образом, предложенная модификация МБП на

несогласованных сетках позволяет получить достаточно точное решение уже на самой грубой сетке.

Заключение. Были реализованы модуль расчета поля давления, основанный на задаче (1)-(3), и модуль постобработки численных потоков. Были проведены вычислительные эксперименты на моделях реальных месторождений. Результаты этих экспериментов показали эффективность МКЭ с методом балансировки потоков по сравнению МКЭ без метода балансировки потоков, а также эффективность предложенной модификации метода балансировки потоков на несогласованных сетках по сравнению с реализацией, предложенной в работе [2].

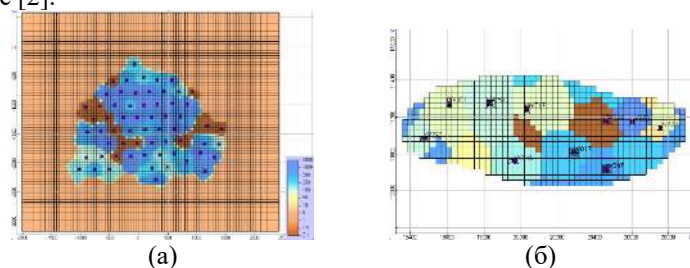


Рис.1 – Распределение структурной проницаемости матрицы–породы на поверхности цифровых двойников M1 (а) и M2 (б)

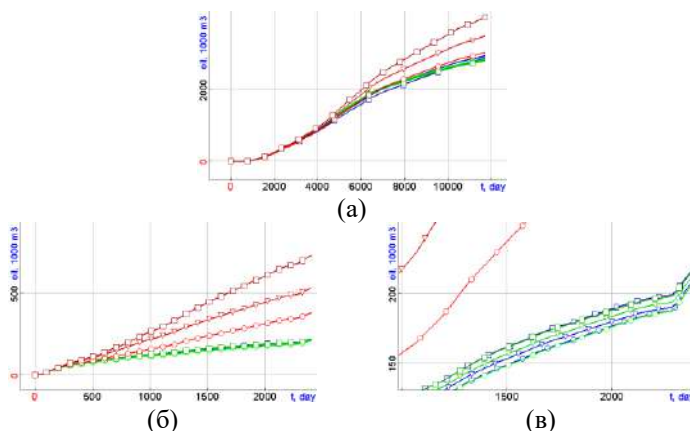


Рис.2 – Графики суммарного отбора нефти для M1 (а) и M2 (б, в) (в – укрупненный масштаб); красный цвет – P1, зеленый цвет – P2, синий цвет – P3; грубой сетке соответствуют квадратики, более подробной – треугольники, самой подробной – кружки

Литература:

1. Соловейчик, Ю.Г. Метод конечных элементов для скалярных и векторных задач / Ю.Г. Соловейчик, М.Э. Рояк, М.Г. Персова. – Новосибирск: НГТУ, 2007. – 869 с.

2. Персова М. Г., Соловейчик Ю. Г., Гриф А. М. Балансировка потоков на неконформных конечноэлементных сетках при моделировании многофазной фильтрации // Программная инженерия. 2021. Т. 12, № 9. С. 450–458.

РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ КОНЕЧНО-ЭЛЕМЕНТНОГО АНАЛИЗА НАПРЯЖЕННО-ДЕФОРМИРОВАННОГО СОСТОЯНИЯ УПРУГИХ ТЕЛ С ПОВРЕЖДЕНИЯМИ

В.С. Слобода

**Новосибирский государственный технический университет,
г. Новосибирск, sloboda1505@gmail.com**

Научный руководитель: Персова М. Г., д. т. н., профессор

Статья посвящена разработке конечно-элементного программного обеспечения для решения двумерных задач теории упругости, в частности, задачи о напряженно-деформированном состоянии панели с круговым отверстием, также известной как задача Кирша. В ходе работы проводится разработка модулей программного обеспечения, валидация моделей, построенных на его основе, а также верификация программного обеспечения на известных аналитических решениях и справочных данных [1].

Ключевые слова: растяжение пластины с отверстием, концентрация напряжений, численное моделирование, метод конечных элементов.

The paper is dedicated to finite element method software development for solving 2D-problems of elasticity theory, in particular, the stress-strain state of a plate with a circular hole problem, also known as Kirsch's problem. During the work, the development of software modules, models built on its basement, and software verification on known solutions and reference books data is considered.

Keywords: tension of a plate with a hole, stress concentration, numerical modeling, finite element method.

В современном машиностроении актуальной и одной из важнейших является проблема предсказания поведения напряжений в окрестностях концентраторов, таких как технологические отверстия и трещины, и их влияния на несущую способность конструкции.

В настоящее время применение зарубежных конечно-элементных пакетов (ANSYS, FEMAP, ABAQUS и др.) для решения этой проблемы

является затруднительным из-за проблем с приобретением лицензий, а также сопряжено с рисками утечек информации об объектах гражданского и военного машиностроения.

Разработанное программное обеспечение (ПО) включает следующие модули:

1. **Модуль обработки геометрии** через консоль запрашивает у пользователя размеры панели, наличие отверстия и его радиус, затем на основе полученных значений создает данные об элементарных подобластях;
2. **Модуль построения конечно-элементной сетки** осуществляет разбиение созданных ранее геометрических примитивов на ячеистую сетку;
3. **Модуль создания конечно-элементной модели** осуществляет реализацию математической модели [2] на полученной сетке, вычисляет локальные матрицы жесткости элементов и векторы нагрузок;
4. **Модуль сборки глобальных матрицы жесткости и вектора нагрузок**;
5. **Модуль решения СЛАУ**;
6. **Модуль постобработки** – осуществляет переход от найденных перемещений к деформациям, затем, в соответствии с законом Гука – к напряжениям внутри элементов.

Программное обеспечение спроектировано в соответствии с принципами объектно-ориентированного подхода (ООП), что позволяет ввести базовый класс элемента и определить его наследников – прямоугольник и четырехугольник. Эти две родственные сущности, несмотря на то, что прямоугольный элемент является частным случаем четырехугольного, и компоненты его локальной матрицы жесткости (МЖ) могут быть получены аналогично четырехугольному элементу, разнесены из-за того, что такой подход увеличивает число вычислительных затрат, связанных с переходом от шаблонного элемента к четырехугольному, а также усложняет масштабирование программного обеспечения. Поэтому при вычислении локальных МЖ проверяется тип элемента и, в зависимости от его типа, программа осуществляет вход в полиморфную функцию, где для прямоугольного элемента проводится интегрирование в естественных координатах, а для четырехугольного – интегрирование в шаблонных координатах с переходом в глобальные через якобиан [3, 4]. Такой подход позволяет избежать лишних вычислений якобианов для прямоугольников и сократить вычислительные затраты при сборке глобальной МЖ.

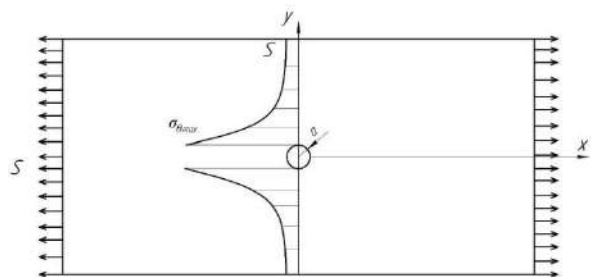


Рисунок 1 – эпюра напряжений в центре панели

Верификация ПО проводится путем сравнения численно полученного решения с аналитическим решением задачи Кирша в виде соотношений

$$\sigma_{\theta \max} = 3S, \quad \sigma_{\theta \min} = -S \quad (1)$$

и эпюры напряжений (рис. 1)

В качестве результата работы ПО принимаются эквивалентные напряжения по критерию Мизеса [5]. Результат такого расчета при приложении единичной нагрузки представлен на рис. 2, сравнение решения, полученного с помощью разработанного ПО, с аналитическим решением, приведено в таблице 1.

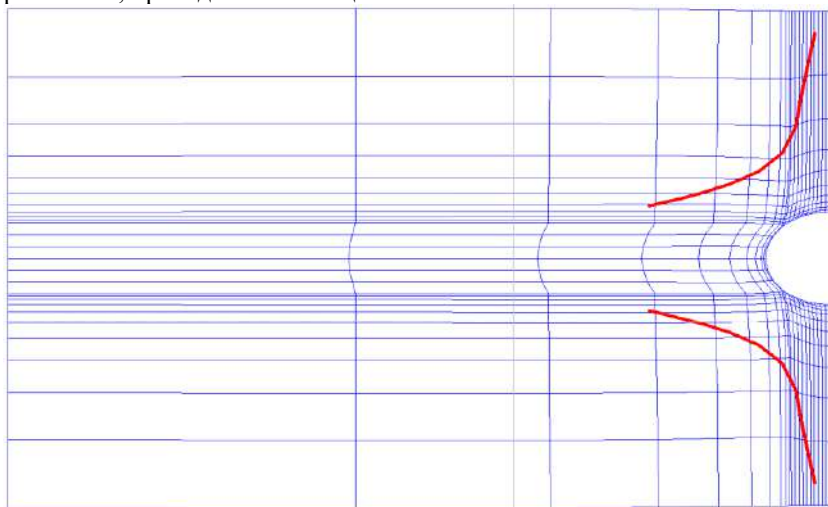


Рисунок 2 – конечно-элементная модель панели и эпюра напряжений, полученные с применением разработанного ПО

Отношение максимальных напряжений к номинальным фактически является коэффициентом концентрации напряжений (ККН)

$$K = \frac{\sigma_{\theta \max}}{\sigma_{\theta \min}} \quad (2)$$

В качестве исследования влияния размера отверстия на величину концентрации приводится график зависимости K_{\max} от отношения диаметра отверстия к ширине панели (рис. 3)

Представленные результаты свидетельствуют об успешной верификации разработанного программного обеспечения. На данном этапе разработки его уже можно использовать для решения простейших задач теории хрупкого разрушения и анализировать несущую способность плоских элементов конструкций. При дальнейшем развитии ПО открываются перспективы проведения комплексных расчетов с более сложной формой, материалами и случаями нагружения.

Таблица 1 – Сравнение результатов работы ПО с аналитическим решением

	$\frac{\sigma_{\theta \max}}{\sigma_{\theta \min}}$	Отклонение Δ
Аналитическое решение	3	0,33%
КЭ решение	2,99	

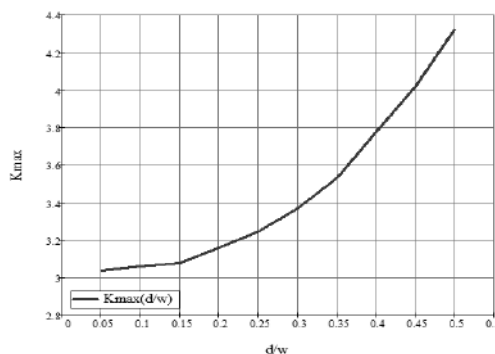


Рисунок 3 – зависимость K_{\max} от размера отверстия

Литература:

1. Howland R.J. On the stresses in the neighborhood of a circular hole in a strip under tension. Ryal Society of London. - Лондон, 1930. – С. 49-86.

2. *Слобода В.С.* Конечноэлементный анализ деформированного состояния упругой панели на регулярной прямоугольной сетке с билинейными базисными функциями // Наука. Технологии. Инновации: XVII Всероссийская научная конференция молодых ученых. 2 ч. – Новосибирск: Изд-во НГТУ. – 2024. – С. 40-43.

3. *Соловейчик Ю.Г., Рояк М.Э., Персова М.Г.* Метод конечных элементов для решения скалярных и векторных задач: учеб. пособие. Новосибирск: Изд-во НГТУ, 2007. – 896 с. («Учебники НГТУ»).

4. *Присекин В.Л., Расторгуев Г.И.* Основы метода конечных элементов: Учеб. пособие. – Новосибирск: Изд-во НГТУ. -2003. – 124 с.

5. *Демидов С.П.* Теория упругости: Учебник для вузов. – М.: Высшая школа, 1979. 432 с.

**Секция ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА,
МАШИННОГО ОБУЧЕНИЯ И ПРЕДСКАЗАТЕЛЬНОГО
МОДЕЛИРОВАНИЯ**

**ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ
В ИНДУСТРИИ 4.0**

Н.А. Бессонов

**Донской государственный технический университет,
г. Ростов-на-дону, 121gug121@gmail.com**

Научный руководитель: Климова Д.Н., к.т.н., доцент

В данной работе исследуется роль технологий машинного обучения и искусственного интеллекта в рамках Индустрии 4.0. Рассмотрены современные интеллектуальные системы управления, позволяющие оптимизировать потребление энергии и адаптировать процессы под изменяющиеся условия эксплуатации. Особое внимание уделено использованию порождающих моделей машинного обучения, способных прогнозировать энергозатраты и адаптировать работу объектов в режиме реального времени. Также анализируются возможности применения Интернета вещей (IoT) для создания «умных» сетей, где автономные объекты обмениваются данными и коллективно повышают энергоэффективность.

This paper examines the role of machine learning and artificial intelligence technologies in the context of Industry 4.0. Modern intelligent control systems that allow optimizing energy consumption and adapting processes to changing operating conditions are considered. Particular attention is paid to the use of generative machine learning models that can predict energy costs and adapt the operation of objects in real time. The possibilities of using the Internet of Things (IoT) to create “smart” networks, where autonomous objects exchange data and collectively improve energy efficiency, are also analyzed.

Цифровая трансформация затрагивает все аспекты бизнеса и корпоративного управления, а применение новых технологий становится ключевым фактором конкурентоспособности компаний. Искусственный интеллект (ИИ) и машинное обучение (МО) представляют собой наиболее важные инструменты в рамках Индустрии 4.0. Эти технологии значительно меняют подходы к управлению, принятию решений и обработке данных, предоставляя компаниям возможность улучшить свои процессы, сократить издержки и повысить эффективность.

Искусственный интеллект и машинное обучение оказывают существенное влияние на бизнес-процессы, внедряясь в такие области,

как прогнозирование спроса, автоматизация производственных линий и улучшение клиентского обслуживания. МО позволяет компаниям анализировать огромные объемы данных, чтобы извлекать ценные инсайты и разрабатывать стратегии на основе данных. Например, с помощью предиктивной аналитики можно прогнозировать изменения рыночных условий и потребительских предпочтений, что помогает бизнесу оставаться гибким и адаптивным в быстро меняющейся среде [1].

Во-первых, современные компании работают в условиях постоянной изменчивости внешней среды. Для того чтобы оставаться конкурентоспособными, бизнес должен адаптироваться к изменяющимся рыночным условиям и эффективно управлять огромными объемами данных. Согласно исследованиям, только 12% компаний, которые внедрили цифровые технологии до пандемии COVID-19, смогли адаптироваться к новым условиям быстро и с минимальными потерями [2]. Это подтверждает необходимость изучения механизмов, способных повысить гибкость и устойчивость бизнеса, таких как ИИ и МО

Вторым важным аспектом является роль больших данных в управлении. Сбор и анализ данных позволяют компаниям принимать более обоснованные и своевременные решения. Тем не менее, объёмы данных продолжают расти, и традиционные системы обработки уже не справляются с их обработкой. Интеграция ИИ и МО в системы больших данных становится критически важной для обеспечения аналитических возможностей компаний. Это позволяет не только улучшить процесс принятия решений, но и выявлять новые рыночные возможности, которые были бы незаметны при традиционных подходах.

Несмотря на все преимущества, которые ИИ, МО предоставляют бизнесу, существуют серьёзные барьеры для их широкого внедрения. Во-первых, многие компании не имеют достаточных ресурсов и знаний для интеграции этих технологий в свои процессы. Во-вторых, отсутствие законодательной базы и регулирования в области ИИ вызывает опасения по поводу этики и правовых вопросов. Например, использование ИИ в принятии решений может привести к проблемам с прозрачностью алгоритмов, а также вызвать риски дискриминации и необоснованных решений.

Особое внимание в рамках исследований ИИ уделяется порождающим моделям в машинном обучении, которые, как показано в статье [3], помогают решать проблемы дефицита данных для обучения моделей. Эти подходы, включая генеративные состязательные сети (GAN) и вариационные автокодировщики (VAE), находят широкое

применение для генерации синтетических данных, что повышает предсказательные способности моделей. Однако встает вопрос о масштабировании этих методов и их применении для сложных и больших наборов данных, что требует дальнейших исследований и доработок.

Оценивание неопределенности в машинном обучении также является важной задачей в контексте корпоративного управления. В работах [4,5] предлагаются методы, такие как кросс-валидация, для улучшения оценки предсказательных способностей моделей, особенно в условиях неопределенности. Это позволяет принимать более точные управленческие решения и минимизировать риски.

ИИ и МО предоставляют компании возможности для автоматизации процессов анализа данных и повышения эффективности прогнозирования. Использование этих технологий позволяет оптимизировать управление ресурсами, снижать издержки и быстро адаптироваться к новым экономическим условиям. Однако их интеграция сталкивается с такими проблемами, как высокая стоимость внедрения, нехватка специалистов и сложности с адаптацией традиционных корпоративных систем под новые требования. Существуют также вызовы, связанные с недостатком законодательного регулирования в области ИИ, что может привести к рискам необоснованных или дискриминационных решений в бизнесе.

Оценивание неопределенности в машинном обучении также является важным элементом для корпоративного управления. Важно не только предсказывать результаты, но и учитывать возможные неопределенности в моделях. Методы, такие как кросс-валидация, помогают повысить точность и достоверность предсказаний, однако внедрение этих подходов в сложные системы, такие как глубокие нейронные сети, также нуждается в дальнейшем развитии.

В контексте цифровой трансформации, ставшей неотъемлемой частью Индустрии 4.0, наблюдается усиление роли сервисной экономики. Принципы клиентоориентированности и использования современных информационно-коммуникационных технологий позволяют компаниям не только повысить эффективность, но и адаптироваться к новым условиям рынка. Важной составляющей сервисной экономики является экономика совместного пользования (шеринговая экономика), которая позволяет потребителям использовать коллективную собственность и тем самым минимизировать личные затраты на покупку товаров и услуг. Этот подход обеспечивает компании доступ к новым сегментам рынка и способствует

устойчивому развитию бизнес-моделей, основанных на принципах цифровизации и совместного потребления.

Кроме того, тенденция к цифровой трансформации связана с внедрением инновационных платформ для анализа и обработки больших данных, что позволяет компаниям выстраивать более гибкие бизнес-модели и быстрее адаптироваться к требованиям рынка. Подход «всё как услуга» (Everything-as-a-Service) и акцент на персонализацию создают новые возможности для повышения уровня клиентского опыта, что становится ключевым фактором конкурентоспособности. Потребительский опыт, ставший центральной целью цифровой трансформации, отражает предпочтение потребителей к удобству и прозрачности. Это требует от компаний интеграции данных и технологий, таких как машинное обучение и искусственный интеллект, для точного предсказания потребностей клиентов и формирования персонализированных предложений [6].

Применение больших данных, машинного обучения и искусственного интеллекта в корпоративном управлении обеспечивает компаниям возможности для более точного анализа и прогнозирования. Анализ больших данных позволяет выявлять скрытые закономерности и тенденции, что способствует принятию более обоснованных решений. Например, машинное обучение может автоматически обрабатывать огромные массивы данных, что снижает трудозатраты и минимизирует ошибки при подготовке отчетности.

Искусственный интеллект также значительно расширяет возможности управления, особенно в таких областях, как управление рисками, прогнозирование финансовых результатов и повышение эффективности производственных процессов. Внедрение ИИ позволяет автоматизировать процессы управления и быстрее адаптироваться к изменениям на рынке, что становится критически важным фактором в условиях глобальной конкуренции.

Интеграция этих передовых технологий в корпоративные процессы не только повышает эффективность управления, но и помогает компаниям укреплять позиции на рынке за счет инноваций и оптимизации внутренних процессов. В то же время, такие изменения требуют от компаний значительных инвестиций в обучение сотрудников и адаптацию корпоративных систем под новые технологии [7].

Помимо преимуществ, Индустрия 4.0 сталкивается с вызовами, связанными с необходимостью надежной защиты данных и обеспечением кибербезопасности. В условиях роста объемов цифровых операций и распространения интернет-устройств возрастает уязвимость

перед кибератаками. Поэтому компании стремятся интегрировать средства кибербезопасности в свои системы, чтобы обеспечить сохранность данных и устойчивость бизнес-процессов.

Литература:

1. *Матвеев А.Н.* Инструменты построения моделей машинного обучения [Электронный ресурс] // E-Scio. – 2023. – №6 (81). – URL: <https://cyberleninka.ru/article/n/instrumenty-postroeniya-modeley-mashinnogo-obucheniya> (дата обращения: 16.09.2024).
2. *Барина Н.В., Барин В.Р.* Цифровая экономика, индустрия 4.0 и искусственный интеллект [Электронный ресурс] // Вестник РЭА им. Г.В. Плеханова. – 2021. – №3 (117). – URL: <https://cyberleninka.ru/article/n/tsifrovaya-ekonomika-industriya-4-0-i-iskusstvennyy-intellekt> (дата обращения: 20.10.2024).
3. *Намиот Д.Е., Ильюшин Е.А.* Порождающие модели в машинном обучении [Электронный ресурс] // International Journal of Open Information Technologies. – 2022. – №7. – URL: <https://cyberleninka.ru/article/n/porozhdayushchie-modeli-v-mashinnom-obucheni> (дата обращения: 16.09.2024).
4. *Арьков В.Ю., Шарипова А.М., Куликов Г.Г.* Оценивание неопределённости в машинном обучении [Электронный ресурс] // Вестник ЮУрГУ. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2023. – №3. – URL: <https://cyberleninka.ru/article/n/otsenivanie-neopredelyonnosti-v-mashinnom-obucheni> (дата обращения: 16.09.2024).
5. *Мейрияни М., Дези Л.В., Видхья А.* Большие данные, машинное обучение, искусственный интеллект и блокчейн в корпоративном управлении [Электронный ресурс] // Форсайт. – 2023. – №4. – URL: <https://cyberleninka.ru/article/n/bolshie-dannye-mashinnoe-obuchenie-iskusstvennyy-intellekt-i-blokcheyn-v-korporativnom-upravlenii> (дата обращения: 20.10.2024).
6. *Блинникова А.В., Данилина О.М., Дашиков А.А.* Цифровая трансформация и шеринговая экономика [Электронный ресурс] // Вестник ГУУ. – 2020. – №8. – URL: <https://cyberleninka.ru/article/n/tsifrovaya-transformatsiya-i-sheringovaya-ekonomika> (дата обращения: 07.11.2024).
7. *Айвазова М.А.* Индустрия 4.0: влияние на экономическое развитие [Электронный ресурс] // Вестник науки. – 2021. – №1 (34). – URL: <https://cyberleninka.ru/article/n/industriya-4-0-vliyanie-na-ekonomicheskoe-razviti> (дата обращения: 07.11.2024).

РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ ОПТИМИЗАЦИИ ПЛАНИРОВАНИЯ ТУРИСТИЧЕСКИХ МАРШРУТОВ

Е.С. Корнева

Новосибирский государственный технический университет,
г. Новосибирск, katerina03535@yandex.ru

Научный руководитель: Эстрайх И.В., ст. преподаватель

В статье приведены результаты разработки приложения для планирования туристических маршрутов. В работе рассматривается применение карт Кохонена для решения задачи кластеризации точек на карте, а также алгоритм Дейкстры для построения маршрутов в полученных кластерах.

The article presents the results of the development of an application for planning tourist routes. The paper considers the use of Kohonen maps to solve the problem of clustering points on the map, as well as Dijkstra's algorithm for building routes in the resulting clusters.

В условиях развития туризма в нашей стране высокую актуальность имеет разработка приложения для оптимизации составления туристических маршрутов с учетом интересов и временных ограничений путешественника. В большинстве случаев туры по городам России не превышают 7 дней, в связи с достаточно сжатыми сроками путешествий есть необходимость разработки программы, которая поможет пользователю составить план поездки для того, чтобы турист смог посетить все желаемые достопримечательности за указанное количество дней.

Туристический маршрут, представляет собой набор точек на карте (достопримечательностей), которые собирается посетить пользователь. Задача планирования и построения маршрута является очень важной и имеет множество вариантов решения.

Для решения оптимизации планирования туристических маршрутов необходимо спроектировать систему, которой смогут воспользоваться не только самостоятельные туристы, но и туристические компании. Система должна решать задачу планирования маршрутов с учетом таких ограничений, как количество дней на прохождение маршрута, время посещения каждой точки, интересы пользователя.

На вход программа получает точки на карте, которые хочет посетить пользователь. В первую очередь для решения задачи необходимо разбить полученные точки на кластеры, для этого используем карты Кохонена.

Карты Кохонена или карты самоорганизации (Self-organizing Map) – это одна из разновидностей нейронных сетей с обучением без учителя, с

помощью которой можно решать различные задачи, в том числе задачу кластеризации и визуализации данных [2][3].

Ключевыми параметрами алгоритма являются:

- Количество нейронов (размер карты) – определяет детализацию карты.
- Количество эпох обучения – предельно допустимое число проходов по набору данных.
- Скорость обучения – скорость, с которой корректируются веса нейронов во время обучения.

На вход алгоритм получает набор точек на карте, каждая точка имеет свои координаты – широту и долготу, а также тип объекта.

При инициализации карты создается сетка нейронов с случайными начальными весами, которые будут обновляться в процессе работы.

На каждом шаге алгоритма обучения необходимо определять лучшее соответствие для выбранного нейрона, для этого вычисляется расстояние между точкой и каждым нейроном на карте, далее выделяется нейрон с минимальным расстоянием.

После чего веса нейрона и его соседей обновляются по следующей формуле:

$$weight_i(t + 1) = weight_i(t) + learningrate(t) * h_i(t) * (x(t) - w_i(t)),$$

где

$weight$ – вектор весов i -го нейрона,

$learningrate(t)$ – скорость обучения на шаге t ,

$h_i(t)$ – функция соседства, определяющая как должны изменяться веса соседних нейронов,

$x(t)$ – входной вектор на шаге t [1].

Процесс обучение повторяется указанное количество эпох, с каждым шагом обеспечивая сходимость алгоритма. После завершения обучения каждой точке сопоставляется нейрон, к которому она ближе, таким образом вокруг полученных нейронов образуются кластеры.

После работы алгоритма получим точки, разбитые на кластеры.

Так, если при инициализации карты Кохонена задать следующие параметры: количество кластеров равное 2, скорость обучение 0.5 и количество эпох равное 1000, используя API Яндекс.Карт, можем вывести на карту следующий результат для точек, расположенных в Новосибирске (Рисунок 1).

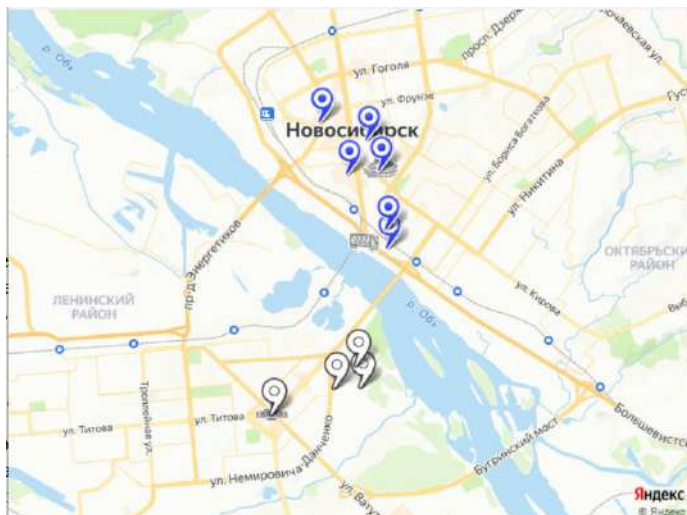


Рис.1 – Результат работы алгоритма кластеризации

После получения списка кластеров, необходимо проложить внутри каждого кластера маршруты. Для этого используем алгоритм Дейкстры, который предназначен для нахождения кратчайшего пути между точками [4].

Каждый кластер представляет собой взвешенный граф. Начальная точка может быть выбрана пользователем или автоматически назначена программой. Далее необходимо найти ближайшую не посещенную точку и добавить ее к маршруту. После чего обновляем позицию и продолжаем составлять маршрут, пока не пройдем все точки кластера. В результате получим – построенный маршрут для каждого кластера.

Приведенный алгоритм дает пользователю возможность гораздо быстрее и удобнее распланировать свою поездку, сформировав оптимальные маршруты на каждый день. Разрабатываемое приложение будет полезно не только самостоятельным туристам, но и туристическим агентствам.

Литература:

1. *Кохонен Т.* Самоорганизующиеся карты. - М.: Бином. Лаборатория знаний, 2008 — 656 с.
2. *Кириченко А.А.* Основы теории искусственных нейронных сетей – М.: Издательская система Ridero, 2020. - 231 с.
3. *Oja E., Kaski S.* Kohonen Maps. – Elsevier, 1999. – 389 с.
4. *Dijkstra E.W.* A note on two problems in connexion with graphs - Numerische Mathematik, 1959, Vol. 1 P. 269–271.

АЛГОРИТМ ПЛАНИРОВАНИЯ СЕАНСОВ ИЗМЕРЕНИЙ ДЛЯ НАВИГАЦИОННЫХ КОСМИЧЕСКИХ АППАРАТОВ

А.А. Кривецкий, В.С. Карманов

Новосибирский государственный технический университет,
г. Новосибирск, kriveczkij.2020@stud.nstu.ru

Рассмотрена задача планирования сеансов измерений между космическими аппаратами глобальной навигационной спутниковой системы, снабженными оптическими терминалами связи. Рассмотрены примеры условий доступности связи и приведен базовый алгоритм планирования расписания сеансов связи.

The problem of scheduling measurement sessions between global navigation satellite system spacecraft equipped with optical communication terminals is considered. Examples of communication availability conditions are considered and a basic algorithm for scheduling of communication session scheduling is given.

Одним из перспективных вариантов развития глобальных навигационных спутниковых систем (ГНСС) является установка аппаратуры для проведения межспутниковых сеансов измерений, использующей оптические терминалы связи. [1]. Эта модификация увеличивает пропускную способность сигнала и увеличивает дальность измерений, способствует синхронизации бортовых шкал времени, что улучшает точность геопозиционирования.

Для планирования таких сеансов связи важно учитывать интервалы взаимной доступности, избегать прохождения оптического луча через атмосферу Земли, где происходит рассеяние, и учитывать засветку приемного терминала солнечным излучением (прямыми или отраженными лучами), которое может помешать передаче данных.

Для вычисления положений космических аппаратов (КА) можно использовать разные методы, например, высокоточные, в которых решается дифференциальное уравнение движения (с использованием данных о параметрах вращения Земли и других возмущающих воздействиях [2-3]) или стандартным методом, на основе решения уравнения Кеплера [4]. В данной работе положения КА моделировались на основе уравнения Кеплера.

Проверка доступности связи

Перед планированием сеансов связи, необходимо составить таблицу взаимной доступности. Таблица является бинарной (принимает значение 0 или 1), индексация по строкам соответствует значению номера пары КА, а индексация по столбцам номеру временного интервала фиксированной длины.

На заданный момент времени, пара КА может осуществлять связь, если выполнены все следующие условия: линия визирования не пересекает атмосферу Земли; расстояние между КА в пределах доступного интервала; угол между линией визирования и направлением на Солнце не меньше заданного значения; угол между линией визирования и направлением на ближайший край диска Земли, освещенного Солнцем не меньше заданного значения; угол между линией визирования и направлением на Луну не меньше заданного значения. Линия визирования – прямая, проходящая между КА, образующими пару.

Проверка всех условий осуществляется методами аналитической геометрии и линейной алгебры [5].

Для проверки условий, необходимо знать положения Луны и Солнца. Для их вычисления можно воспользоваться различными методами [6]. В данной работе используется модель движения DE421 [7]. Данные предоставляются онлайн службой эфемерид Института прикладной астрономии РАН [8].

Алгоритм планирования сеансов связи

Задача планирования сеансов связи межспутниковых измерений, сводится к задаче оптимизации на перестановках (задача составления расписаний). Задача является NP-трудной, что исключает использование методов полного перебора. Предлагается использование жадного алгоритма и разбиение всего промежутка планирования на интервалы синхронизации шкал времени (СШВ) равной длины $\Delta_{\text{СШВ}}$.

Задача планирования может иметь различные постановки, например:

1. $N \rightarrow \max, T = \text{const}$
2. $N = \text{const}, T \rightarrow \min,$
3. $N \rightarrow \max, T \rightarrow \min,$

где N – количество пар КА-КА, выполнивших связь на интервале планирования, T – время за которое эти сеансы выполняются.

На каждом КА установлено M терминалов двусторонней связи (одновременный прием и передача для пары КА).

Входные данные алгоритма:

- Матрица доступности сеансов связи: V ,
- Количество КА: n ,
- Количество терминалов связи на одном КА: M ,
- Время необходимое для выполнения сеанса связи t .

Выходные данные:

Оптимальный план межспутниковых сеансов измерений.

Рассмотрим общую схему алгоритма. На каждом шаге размещать сеанс на интервале с минимальным значением функции веса $w_{i,j}^l$. Такой

подход позволит получать субоптимальные планы сеансов. Изменяя функцию веса, можно достигать разных характеристик плана.

Алгоритм планирования:

1. Вычислить списки взаимной доступности связи l для каждой пары КА.
2. Удалить (сделать недоступными для связи) интервалы доступности, которые короче времени t .
3. Повторить 1. Переход на 4.
4. Для каждой пары i, j и каждого интервала доступности l вычислить вес $w_{i,j}^l$.
5. Расположить сеанс связи в начале интервала с минимальным весом.
6. Пересчитать интервалы видимости.
7. Пока не останется интервалов видимости, повторять 1, иначе план составлен.

Примеры функций веса:

- 1) $w_{i,j}^l = (t_{i,j}^{l,0} + 1)(t_{i,j}^{l,1} - t_{i,j}^{l,0})(c_{i,j} + 0,1)$, направлена на создание плана на больших временных промежутках, максимизируется количество связей, минимизируется время.
 - 2) $w_{i,j}^l = (t_{i,j}^{l,1} - t_{i,j}^{l,0})(c_{i,j} + 0,1)$, направлена на создание плана на небольших временных промежутках, максимизируется количество связей, минимизируется время.
 - 3) $w_{i,j}^l = c_{i,j}$, направлена на поиск минимального времени, за которое вся группировка может совершить сеансы связи,
- где $t_{i,j}^{l,0}$ – индекс начала интервала l для пары КА (i, j) , $t_{i,j}^{l,0}$ – индекс конца интервала l для пары КА (i, j) , $c_{i,j}$ – счетчик, сохраняющий количество проведенных сеансов для пары КА (i, j) .

Численный эксперимент

Рассмотрим вычисленный фрагмент плана. Входные параметры:

Шаг по времени: 10 с;

Длина интервала: $\Delta_{\text{СШВ}} = 3300$ с;

Количество интервалов $\Delta_{\text{СШВ}}$: 5;

Функция веса: $w_{i,j}^l = (t_{i,j}^{l,0} + 1)(t_{i,j}^{l,1} - t_{i,j}^{l,0})(c_{i,j} + 0,1)$;

Диаграмма Ганта для первого интервала $\Delta_{\text{СШВ}}$ представлена на рисунке 1. Рассмотрены все пары, образуемые КА №8. Черным цветом помечены интервалы, в которых связь недоступна, красным – этап подготовки к сеансу, зеленым сам сеанс измерений.

Полученная диаграмма, отображает план, в котором максимизируется количество связей за фиксированное время.

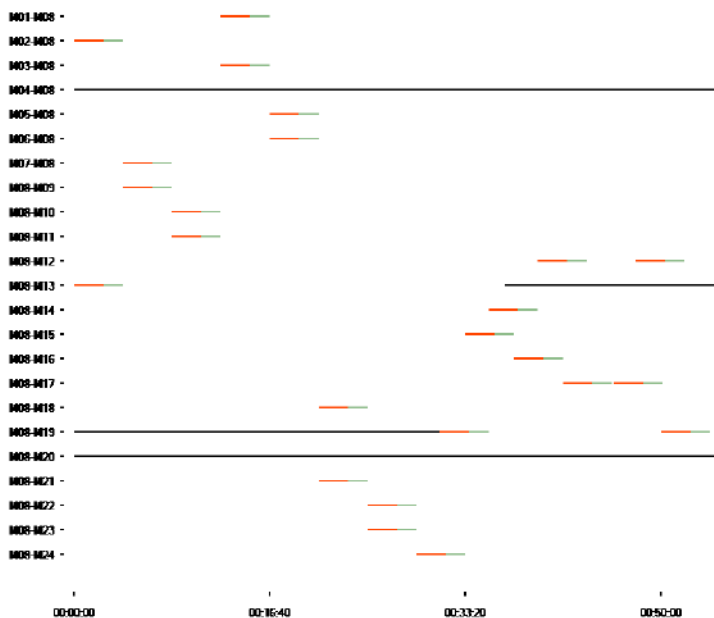


Рис. 1 – Диаграмма Ганта составленного плана. Первый интервал $\Delta_{\text{СШВ}}$, КА №8.

Заключение

Разработанный подход может найти применение в перспективных системах глобальной спутниковой навигации. Приведенный алгоритм планируется исследовать более детально и сравнивать с другими (например, стохастическими).

Литература:

1. *Бартенев В.А. [и др.].* Современные и перспективные информационные ГНСС - технологии в задачах высокоточной навигации / ред. В.А. Бартенев, М.Н. Красильщиков. – М.: Физматлит, 2014. – 192 с.
2. *Гречкосеев А.К., Толстиков А.С., Тиссен В.М., Карманов В.С., Ваганова А.И.* Модификация базового метода сингулярного спектрального анализа для повышения точности прогнозирования неравномерности вращения Земли // *Вычислительные технологии.* – 2020. – Т. 25, № 3. – С. 54–65.
3. *Кривецкий, А.А.* Сравнение алгоритмов прогноза в задаче определения неравномерности вращения земли / А. А. Кривецкий //

Наука. Технологии. Инновации: XV Всероссийская научная конференция молодых ученых, посвященная Году науки и технологий в России. Сборник научных трудов в 10-ти частях, Новосибирск, 06–10 декабря 2021 года. Том Часть 2. – Новосибирск: НГТУ, 2021. – С. 194–198.

4. *Кривецкий, А.А.* Моделирование эфемерид перспективных космических аппаратов / А. А. Кривецкий // Наука. Технологии. Инновации: Сборник научных трудов XVII Всероссийской научной конференции молодых ученых. В 11-ти частях, Новосибирск, 04–08 декабря 2023 года. – Новосибирск: НГТУ, 2024. – С. 62–66.

5. *Денисов В.И.* Алгебра и геометрия. Практикум: [учебник] / В.И. Денисов, В.М. Чубич, О.С. Черникова. – Новосибирск: Изд-во НГТУ, 2018. – 305 с.

7. *Абалакин В.К., Аксёнов Е.П., Гребеников Е.А., Дёмин В.Г., Рябов Ю.А.* Справочное руководство по небесной механике и астродинамике. Под ред. Г.Н. Дубошина. М., Наука, 1976.

6. *Folkner W., Williams J., Boggs D.* The planetary and lunar ephemeris DE 421 [Electronic resource] // IPN Progress Report 42-178. – 2009. – Vol. 005. – P. 1–34. – URL: <http://tmo.jpl.nasa.gov/progress/report/42-178/178C.pdf>.

7. Онлайн служба эфемерид ИПА РАН [Электронный ресурс]. – URL: <https://iaaras.ru/dept/ephemeris/online/> дата обращения: 20.05.2024.

ПРОГНОСТИЧЕСКАЯ МОДЕЛЬ ДЛЯ ОЦЕНКИ ЭФФЕКТИВНОСТИ ПРОГНОЗИРОВАНИЯ АВТОСТРАХОВЫХ СЛУЧАЕВ

К.Г. Мадияров

Новосибирский государственный университет

экономики и управления «НИНХ»,

г. Новосибирск, kuan.mad@mail.ru

Научный руководитель: Пестунов А.И., к. физ.-мат. н., доцент

В сложных климатических условиях, встречающихся на территории Казахстана особую значимость приобретает прогнозирование страховых случаев при ненадлежащем вождении (на примере снежных буранов). В данной статье описана методология предиктивной модели с помощью методов ML. Проводится оценка влияния ненадлежащего вождения и погодных рисков (на примере снежных буранов) на вероятность наступления ДТП. Рассматриваются аспекты использования прогностической модели на основе традиционных актуарных методов и методов, основанных на ML.

Осуществлена проверка повышения точности прогнозирования с помощью методов ML в сравнении с актуарными методами.

In the difficult climatic conditions encountered in Kazakhstan, forecasting insurance cases due to improper driving (using snowstorms as an example) is of particular importance. This article describes the methodology of a predictive model using ML methods. An assessment is made of the impact of improper driving and weather risks (using snowstorms as an example) on the probability of an accident. Aspects of using a predictive model based on traditional actuarial methods and ML-based methods are considered. An increase in forecasting accuracy using ML methods in comparison with actuarial methods is verified.

Страховые компании сталкиваются с необходимостью прогнозирования страховых случаев с учетом погодных рисков. В исследованиях авторов [6] отражено влияние изменения климата на страховой сектор. Традиционные актуарные методы не всегда помогают провести точный анализ, т.к. не представляется возможным работать с динамическими данными. Решением данной проблемы могут служить методы ML. В исследованиях авторов [4], [7] модели CNN позволяют оценивать поведение водителей. Маглинова Т.Г. и Шупило О.М [3] предлагают внедрение AI в страховые компании для повышения точности моделирования рисков.

Для тестирования метода использованы данные, собранные из исторических записей страховых компаний, Базы данных гидрометцентров и статистических отчетов.

Таблица 1 – Показатели актуарных расчетов

Метеоданные	Сред. ск-ть ветра	Инт. снего-пада	Видимость	Прод. бури
	14,6 м/с	1,8 см/час	0,47 км	5,5 ч.
Поведение	Сред. ск-ть вод-ля	Част. торможения	Уровень аварийности	
	57,3	27 раз/час	1 авария на 18 вод.-ей	
Претензии	Сред. сумма страхового возм. По несч. случ., связ. с бурей		Частота претензий на 100 вод-й во вр. бурь	
	2,370,000		5,6	

Показатели актуарных расчетов приведены в таблице 2. В данном случае показатель SSI – это индекс интенсивности снежных бурь. Показатель высокого погодного риска.

Таблица 2 – Показатели актуарных расчетов

Частота	Тяжесть	Ожидаемая стоимость	Поправка с учетом SSI	Премия
5,6	473,000	13,272,000	18,484,326	22,181,192

Данные для прогностических моделей ML представлены в таблице 3. RFD – это фактор риска поведения водителя. Служит показателем рискованного поведения при вождении. Скорректированная средняя скорость служит индикатором осторожности водителей. Если этот показатель высокий, то есть вероятность, что водитель относится пренебрежительно к погодным рискам. Чем ниже данный показатель, тем лучше. Это означает что водитель не демонстрирует рискованное поведение.

Таблица 3 – Объединенные данные модели ML

SSI	RFD	Скорректир. ср. ск-ть
5,6	2,12	1,03

Результаты. Оценка характеристик эффективности прогностических моделей, основанные на ML и включающие библиотеки логистической регрессии, метод случайного леса, XGBoost приведены в таблице 4.

Таблица 4 – Показатели эффективности моделей

Модель	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Логистическая регрессия	0,78	0,72	0,75	0,73	0,76
Случайного леса	0,84	0,79	0,82	0,80	0,85
XGBoost	0,87	0,81	0,85	0,83	0,88

На рисунке 1 приведенной диаграмме отражен сравнительный анализ эффективности актуарных методов и методов ML. Согласно рисунку, можно увидеть, что машинные прогностические модели точнее, чем традиционные.

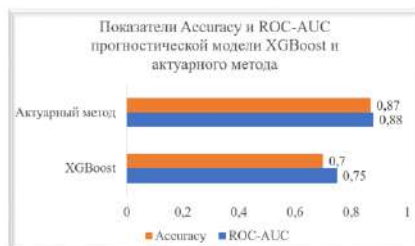


Рис.1 – Сравнение показателей Accuracy и ROC-AUC прогностической модели XGBoost и актуарного метода.

Обсуждение результатов. При интеграции данных об интенсивности снежных бурь и данных о поведении водителей представляется комплексная структура оценки риска. Это подтверждает исследования Васиной Д. И. [1] и Плоткиной А. А. [2] о положительном эффекте внедрения ML при прогнозировании страховых случаев. Также исследования Mosimann M., [5] и др., Conforti R., [8] и др., о том, что модели ML имеют множество преимуществ.

Заключение. Эффективность прогнозирования страховых случаев возможно повысить с помощью ML. Прогностическая модель повышает точность прогнозирования на основе гидрометеорологических данных и поведении водителя за рулём может оказать помощь страховым компаниям при принятии решений о корректировке страховых премий наряду с традиционными актуарными методами.

Литература:

1. *Васина Д.И.* Обзор и апробирование методов машинного обучения для краткосрочного прогнозирования // Математические структуры и моделирование. – 2023. – № 4 (68). – С. 59-70. – doi: 10.24147/2222-8772.2023.4.59-70.
2. *Плоткина А.А.* Использование методов машинного обучения в имущественном страховании // Высокие цифровые технологии: от анализа данных к искусственному интеллекту, 2021. – С. 260-263.
3. *Маглинова Т.Г., Шупило О.М.* Внедрение искусственного интеллекта в страховую отрасль // Международный журнал гуманитарных и естественных наук. – 2022. – № 5-4 (68). – С. 142-145. – doi:10.24412/2500-1000-2022-5-4-142-145.
4. *Varaich Z.A., Khalid S.* Recognizing Actions of Distracted Drivers using Inception v3 and Xception Convolutional Neural Networks // Proceedings of the 2019 International Conference on Advanced Computing and Communication Systems (ICACS). – 2019. – P. 347-351. – doi: 10.23919/ICACS.2019.8689131.

5. Mosimann M., Kauzlaric M., Schick S., Martius O., Zischg A.P. Evaluation of surrogate flood models for the use in impact-based flood warning systems at national scale // Environmental Modeling & Software. – 2023. – Vol. 147. – doi: 10.1016/j.envsoft.2023.105936.
6. Zobaа, A. F. Climate Change Risks and Financial Sector // 2005 IEEE Power Engineering Society General Meeting, June 16, 2005. IEEE, 2005. – doi: 10.1109/PES.2005
7. Li Y., Sun C., Hu Y. Whale Optimization Algorithm-Based Deep Learning Model for Driver Identification in Intelligent Transport Systems // Computers, Materials & Continua. – 2023. – Vol. 77. – P. 98-106. – doi: 10.32604/cmc.2023.035878.
8. Conforti R., Leoni M.D., La Rosa M., Aalst W.M.P.V.D., Hofstede A.H.M.T. A recommendation system for predicting risks across multiple business process instances // Decision Support Systems. – 2014. – Vol. 67. – P. 89-102. – doi: 10.1016/j.dss.2014.10.006.

ФИЛЬТРАЦИЯ ШУМА В ЭТНОГРАФИЧЕСКИХ АУДИОЗАПИСЯХ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Л.А. Миронов

**Новосибирский государственный технический университет,
г. Новосибирск, l.mironov.2020@stud.nstu.ru**

Научный руководитель: Попов А. А., д.т.н., профессор

В статье рассматриваются методы фильтрации шума в этнографических аудиозаписях фольклорной музыки. Проведено сравнение традиционных методов (вейвлеты, NMF) и нейросетевого подхода на основе генеративной модели Demucs.

This paper explores noise reduction techniques for ethnographic audio recordings of folk music. Traditional methods (wavelet, NMF) are compared with a neural network approach based on the Demucs generative model.

Этнографические аудиозаписи, являющиеся ценным источником информации о культурном наследии, часто страдают от низкого качества звука, вызванного различными типами шумов: фоновым шумом, помехами от оборудования, артефактами хранения носителей и др. Эти шумы значительно затрудняют анализ и интерпретацию записей, снижая их научную и культурную ценность. Существующие методы шумоподавления не всегда обеспечивают достаточную эффективность при сохранении важных акустических характеристик этнографических записей, особенно в случаях сложных и

нестационарных шумов. Таким образом, существует проблема разработки эффективных методов фильтрации шума в этнографических аудиозаписях, которые позволят улучшить качество звука, сохраняя при этом аутентичность и информативность записей.

В работе были рассмотрены некоторые традиционные методы фильтрации. Методы на основе вейвлетов с использованием кратномасштабного анализа [1], позволяющие анализировать сигнал с разным временным и частотным разрешением, метод неотрицательной матричной факторизации (NMF), разделяющий спектрограмму сигнала на компоненты, соответствующие чистому сигналу и шуму [2]. Реализация и тестирование традиционных методов были проведены на основе библиотек PyWavelets и Librosa. Метод на основе вейвлетов с использованием soft thresholding показал удовлетворительные результаты, однако метод NMF, несмотря на более высокое значение SNR после фильтрации, внес нежелательные артефакты в сигнал.

В качестве основного нейросетевого подхода была выбрана генеративная модель Demucs [3] (схема архитектуры представлена на рисунке 1), изначально разработанная для разделения источников звука, основана на каскаде кодировщиков и декодировщиков со skip-connections и LSTM слоем в латентном пространстве. Данная архитектура использует волновую форму сигнала вместо спектрограммы, что позволяет избежать наложению дополнительных шумов из фазы.

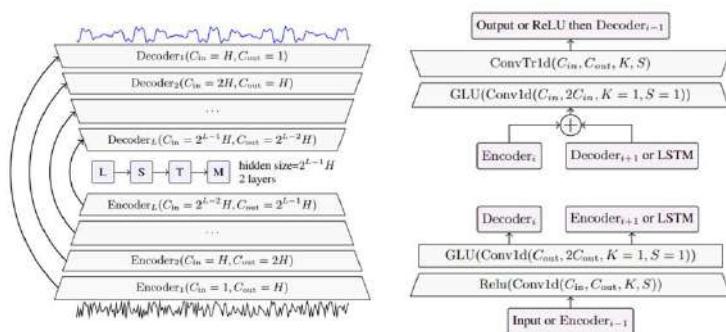


Рис.1 – схема архитектуры Demucs [3]

Для обучения нейронной сети был собран набор синтетически сгенерированных данных из пар $\langle \text{clean}, \text{noisy} \rangle$, где clean – чистый сигнал, noisy – зашумленная версия этого сигнала. В качестве чистого сигнала были использованы музыкальные фольклорные композиции, записанные в современных условиях, без посторонних шумов. Так же

для увеличения объема данных была использована аугментация, включающая в себя изменение высоты тона, полярности, громкости, временная и частотная маскировка [4]. Зашумленная версия сигнала получается путём добавления шумов, соответствующим шумам аппаратуры, внешним шумам, дефектам записи и т. д.

В качестве функции потерь использовалась комбинация L1 нормы и STFT-функцией потерь по нескольким разрешениям (Multiresolution STFT loss [3]), учитывающая спектральные и амплитудные характеристики сигнала в разных частотных диапазонах:

$$L^{(i)}(y, \hat{y}) = L_{sc}(y, \hat{y}) + L_{mag}(y, \hat{y}),$$

$$L_{sc}(y, \hat{y}) = \frac{|STFT(y) - STFT(\hat{y})|_F}{|STFT(y)|_F},$$

$$L_{mag}(y, \hat{y}) = \frac{1}{T} |\log STFT(y) - \log STGT(\hat{y})|_F,$$

где $STFT$ – действительная часть кратковременного преобразования Фурье T – длина сигнала y .

Для оценки качества отфильтрованных записей использовались метрики SNR (Signal to Noise Ratio), RSQI (Relative Signal Quality Improvement), как отношение приращения SNR по отношению к SNR загрязненного сигнала, и deltaVGG, основанная на сравнении эмбедингов предобученной сети VGG16 [5].

В результате ряда исследований были выбраны следующие параметры модели: количество слоёв – 5, частота дискретизация входного сигнала – 22050 Гц, начальная скорость обучения – $4e-4$, длина входного слайса – 2 с. Итоговое обучение модели проводилось с использованием графического ускорителя. Размер батча – 90. В результате чего удалось добиться функции потерь в пределах $2e-1$, среднего значения RSQI около 1.5 и метрики deltaVGG около $1e4$.

Для сравнения были выбраны две аудиозаписи: этнографическая и синтетическая. Результаты работы методов для этнографической аудиозаписи представлены на рисунке 2, а для синтетической – на рисунке 3.

Здесь изначальная Demucs реализация – реализация данной архитектурой для задачи фильтрации шума в аудиозаписях с речью.

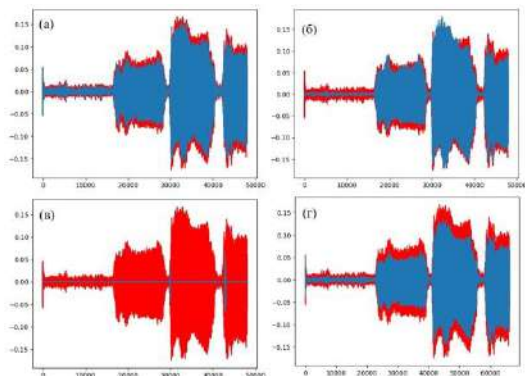


Рис.2 – результат работы методов, (а) – на основе вейвлетов, (б) – на основе NMF, (в) – изначальная Demucs реализация, (г) – итоговая модель (красным цветом обозначен зашумленный сигнал, синим – отфильтрованный)

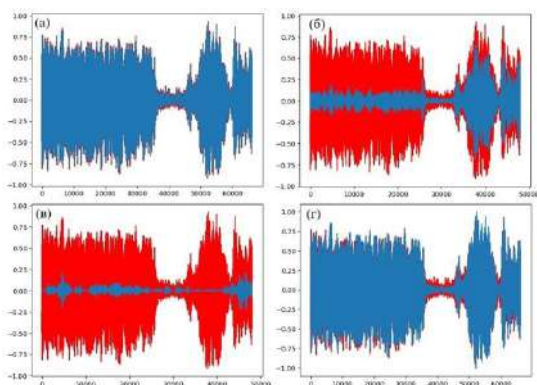


Рис.3 – результат работы методов, (а) – на основе вейвлетов, (б) – на основе NMF, (в) – изначальная Demucs реализация, (г) – итоговая модель (красным цветом обозначен зашумленный сигнал, синим – отфильтрованный)

Также результаты для обеих аудиозаписей были проанализированы с помощью метрики SNR (таблица 1).

Результаты исследования показали, что нейросетевой подход, основанный на генеративной модели Demucs, обладает значительным потенциалом для решения задачи шумоподавления в этнографических аудиозаписях. Такой подход, хоть и не всегда показывает значительные результаты по объективным метрикам, выигрывает в субъективном

восприятию аудиозаписей. Это ставит перед нами проблему разработки дополнительных метрик для оценки отфильтрованных сигналов (например, психоакустических).

Таблица 1 – Сравнение методов фильтрации шума

Запись	Этнографическая запись			Синтетическая запись		
Показатель	SNR_1 SNR_1 , дБ	SNR_2 SNR_1 , дБ	ΔSNR ΔSNR SNR_1 , %	SNR_1 SNR_1 , дБ	SNR_2 SNR_1 , дБ	ΔSNR SNR_1 , %
Wavelet	17.77	19.70	10.88	7.84	8.79	12.13
NMF		24.16	35.99		9.37	18.39
Demucs(f)		16.88	-5.27		8.43	7.46
Demucs		20.54	15.30		10.50	33.88

Литература:

1. Айфичер Э.С. Цифровая обработка сигналов: практический подход / Э. С. Айфичер, Б. У. Джервис. – 2-е изд. – М.: Издательский дом «Вильямс», 2004. – 992 с.
2. Iwase Y., Kitamura D. Supervised Audio Source Separation Based on Nonnegative Matrix Factorization with Cosine Similarity Penalty // IEICE trans. fundamentals. – 2015. – Vol.E98-A, № 1. – 1-8 p.
3. Defossez, A. Real Time Speech Enhancement in the Waveform Domain / A. Defossez, G. Synnaeve, Y. Adi // Facebook AI Research INRIA PSL Research University. – 2020. – 5 p.
4. Park, D.S. SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition / D. S. Park, W. Chan, Y. Zhang [и др.] // Google Brain, 2019. – 6 p.
5. Li, Y. Learning to denoise historical music / Y. Li, B. Gfeller, M. Tagliasacchi [и др.] // Google, 2022. – 8 p.

СПОСОБНЫ ЛИ СОВРЕМЕННЫЕ СЕРВИСЫ, ОСНОВАННЫЕ НА БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЯХ, ДАВАТЬ ОЦЕНКИ ПРОТОТИПАМ ДИЗАЙНОВ?

Д.А. Попова, Ф.А. Ткачук

Новосибирский государственный технический университет,

г. Новосибирск, darya.sam54@gmail.com

Научный руководитель: Бакаев М.А., к.т.н., доцент

Статья дает анализ возможностей современных сервисов на базе больших языковых моделей (LLM) для автоматизированной оценки прототипов дизайна веб-сайтов. В работе подробно освещается сравнительное исследование оценок, выданных LLM-сервисами Rate My Image (RMI) и ChatGPT, и экспертами-дизайнерами. Анализ выявил разницу, что RMI, фокусируясь на эстетике, завышает оценки на 39% по сравнению с экспертами, в то время как ChatGPT, оценивая детализацию и функциональность, демонстрирует меньшее расхождение. Это обусловлено как субъективностью экспертных оценок, так и спецификой алгоритмов ИИ. Результаты исследования позволяют определить сильные и слабые стороны применения LLM для оценки дизайна и указывают на необходимость дальнейших исследований для повышения объективности автоматизированной оценки.

The paper analyses the capabilities of modern services based on large language models (LLM) for automated evaluation of website design prototypes. The paper details a comparative study of evaluations given by the LLM services Rate My Image (RMI) and ChatGPT, and by expert designers. The analysis reveals a difference that RMI, focusing on aesthetics, overestimates by 39% compared to experts, while ChatGPT, evaluating detail and functionality, shows a smaller discrepancy. This is due to both the subjectivity of expert evaluations and the specificity of AI algorithms. The results of the study identify the strengths and weaknesses of applying LLM to design evaluation and indicate the need for further research to improve the objectivity of automated evaluation.

Введение

Оценка пользовательского интерфейса (UI) – сложная и трудоемкая задача, критически важная для разработки программного обеспечения. Хотя генеративные модели AI/ML предлагают потенциал для автоматизации этого процесса, оценка UI на основе прототипов дизайна сайта всё ещё находится на стадии развития. Современные сервисы на базе больших языковых моделей (LLM) предлагают быстрый анализ, но их способность давать объективные и всесторонние оценки требует тщательного изучения [1].

Автоматизация оценки дизайна веб-сайтов с помощью искусственного интеллекта (далее по тексту ИИ) привлекает все больше внимания. Современные сервисы, основанные на больших языковых моделях (LLM), предлагают потенциально быстрый и эффективный способ анализа прототипов дизайна [2]. Однако, способность LLM предоставлять объективные и полные оценки остается предметом исследования [3]. Данная работа посвящена изучению возможностей современных сервисов на основе LLM в области оценки прототипов дизайна веб-сайтов, анализируя их сильные и слабые стороны и определяя границы их применимости.

Исходя из этой мотивации, в данной статье рассматривается следующий исследовательский вопрос: "Возможно ли использовать ИИ

для оценки прототипов дизайнов и могут ли ИИ заменить оценки дизайнеров на этапах проектирования?".

В основной части оценивались прототипы дизайнов сайтов двумя ИИ: Rate My Image (далее RMI) и ChatGPT - по десятибалльной шкале, также были даны пояснения к оценкам. Было обращено внимание на то, что RMI при оценке уделяет большое внимание эстетичности (критерий A), а ChatGPT – уровню детализации (критерий D). Далее сравнивались оценки экспертов из статьи [4], и ИИ по аналогичным критериям, чтобы определить соответствие и выявить области успешного применения и ограничения ИИ в автоматизированной оценке дизайна.

Основная часть

Оценивались 84 изображения сгенерированных при помощи ИИ, способных генерировать изображение по текстовому запросу (KreaAI, Stable Diffusion и LoRA) [5]. С помощью сервисов RMI (визуальная составляющая) и ChatGPT (информационная и функциональная составляющие), путём просьбы оценить изображение по десятибалльной шкале и дать пояснение оценке. В ходе оценки были выявлены ключевые факторы, определяющие оценку. Сервис RMI показал значительное влияние визуального оформления, где эстетика и современность, качество и интеграция изображений, цветовая гамма и композиционная структура играли важную роль. Анализ данных ChatGPT подчеркнул важность информативности и юзабилити, выделив ясность цели, качество контента, навигацию, интерактивность, функциональность и мобильную адаптацию как важные аспекты. Примечательно, что низкое качество типографики оказывало сильное негативное влияние на общую оценку.

Для сравнительного анализа оценок, полученных от дизайнеров-экспертов и от автоматизированных систем, было принято решение сопоставить оценки эстетической составляющей дизайна (критерий A) от дизайнеров с оценками, предоставленными сервисом RMI. Выбор данного сервиса обусловлен ориентацией его критериев оценки на визуальную привлекательность дизайна. Аналогично, для оценки уровня детализации (критерий D) использовались данные сервиса ChatGPT, учитывая, что значительная часть его критериев оценки основана на детализации и качестве контентного наполнения.

Анализ данных, представленных в таблице 1, показывает наименьшее расхождение между оценками, выданными моделью ChatGPT, и экспертными оценками по критерию D (уровень детализации). В то же время, наблюдается существенное расхождение оценок RMI по критерию A (эстетика дизайна) с экспертными.

Оценки RMI в среднем на 39% превысили экспертные оценки. Несмотря на инструкцию оценивать исключительно по критерию А, в экспертных оценках, вероятно, присутствует влияние человеческого фактора, включая субъективные восприятия, помимо чистого критерия привлекательности. В то же время, алгоритм RMI, хотя и нацелен на оценку привлекательности, может использовать в своей работе другие, неявно заданные параметры, влияющие на итоговый результат и приводящие к завышенным оценкам по сравнению с экспертными.

Таблица 1 – Результаты оценок прототипов дизайнов

Сервис оценки	Запрос			
	Long, with image	Long, no image	Short, with image	Short, no image
ChatGPT	KreaAI: 4.3 LoRA: 2.3 SD 1.5: 3.3	KreaAI: 4 LoRA: 3 SD 1.5: 5.3	KreaAI: 2.6 LoRA: 4 SD 1.5: 2	KreaAI: 4 LoRA: 4.3 SD 1.5: 3
Эксперты (D)	KreaAI: 6.14 LoRA: 2.86 SD 1.5: 4.21	KreaAI: 5.00 LoRA: 2.92 SD 1.5: 3.14	KreaAI: 6.07 LoRA: 2.79 SD 1.5: 4.58	KreaAI: 4.93 LoRA: 2.92 SD 1.5: 3.21
RMI	KreaAI: 4 LoRA: 3 SD 1.5: 4	KreaAI: 6 LoRA: 5 SD 1.5: 4	KreaAI: 7 LoRA: 6 SD 1.5: 5	KreaAI: 8 LoRA: 4 SD 1.5: 6
Эксперты (A)	KreaAI: 4.85 LoRA: 3.07 SD 1.5: 3.64	KreaAI: 4.14 LoRA: 3.21 SD 1.5: 3.21	KreaAI: 5.28 LoRA: 3.21 SD 1.5: 3.92	KreaAI: 4.50 LoRA: 2.92 SD 1.5: 3.00

Сравнительный анализ экспертных оценок и моделью ChatGPT по критерию D (уровень детализации), показал статистически значимое различие. Оценки, выданные моделью ChatGPT, в среднем на 7% ниже экспертных оценок. Это различие, предположительно, обусловлено меньшим весом, придаваемым моделью ChatGPT эстетическим аспектам дизайна, в то время как эксперты, оценивая уровень детализации, учитывают эти факторы в комплексе с другими параметрами, влияющими на восприятие детализации.

Заключение

Результаты исследования демонстрируют перспективность использования модели ChatGPT для автоматизированной оценки уровня детализации веб-сайтов. Незначительное расхождение (7%) оценок ChatGPT с экспертными оценками, в отличие от существенного расхождения с оценками RMI (39%), подтверждает эту возможность.

Низкий уровень расхождений позволяет предположить применимость ChatGPT и для оценки функциональности сайтов, предполагая корреляцию между этими параметрами. Однако, анализ сопроводительных текстов выявил влияние качества текстового контента прототипов на оценки ChatGPT. Рекомендуется использовать прототипы с полноценным текстовым контентом для повышения точности оценки.

Литература:

1. Yang, B., & Li, S. (2024). UISGPT: Automated Mobile UI Design Smell Detection with Large Language Models. Electronics. – Vol. 13(16), – P. 3127.
2. Костинич К.В., Видьманов Д.А. Мобильный пользовательский интерфейс на основе генеративных моделей. // Материалы XV Всероссийской научно-технической конференции с международным участием «робототехника и искусственный интеллект», 2023. – С. 136-141.
3. Takaffoli, M., Li, S., & Mäkelä, V. (2024, July). Generative AI in User Experience Design and Research: How Do UX Practitioners, Teams, and Companies Use GenAI in Industry?. In Proceedings of the 2024 ACM Designing Interactive Systems Conference, 2024. – P. 1579-1593.
4. Vologdin D., Bakaev M. (2024) Multimodal Prompts in AI-Assisted User Interface Design: How Many Words a Picture is Worth? Int. Conf. Internet and Modern Society (IMS-2024), St. Petersburg. - в печати.
5. Feng, S., Yuan, M., Chen, J., Xing, Z., & Chen, C. (2023). Designing with Language: Wireframing UI Design Intent with Generative Large Language Models. arXiv preprint arXiv:2312.07755.

**АНСАМБЛЕВАЯ МОДЕЛЬ ГЛУБОКОГО ОБУЧЕНИЯ
ДЛЯ ОБНАРУЖЕНИЯ МОШЕННИЧЕСКИХ ДЕЙСТВИЙ
В ФИНАНСОВЫХ НАБОРАХ ДАННЫХ**

Р.В. Самойленко

Новосибирский государственный университет

экономики и управления «НИНХ»,

г. Новосибирск, ro999@yandex.ru

Научный руководитель: Пестунов А.И., к. ф-м. н., доцент

Статья посвящена исследованию проблем противодействия мошенничеству в сфере электронной коммерции, посредством использования алгоритмов глубокого обучения. Данный материал может быть полезен при

изучении теории анализа данных, а также использован финансовыми организациями для улучшения используемых методик обнаружения недобросовестных денежных транзакций.

The article is devoted to the study of the problems of countering fraud in the field of e-commerce, through the use of deep learning algorithms. This material can be useful in studying the theory of data analysis, as well as used by financial organizations to improve the methods used to detect unfair monetary transactions.

Цифровые технологии произвели революцию в финансовой сфере. В наше время жизнь без интернет-банкинга, онлайн-торговли и мобильных платежных систем практически немыслима. Это приносит удобство, но также открывает новые горизонты для мошенников. Они подстраиваются под современные технологии и создают всё более хитроумные схемы.

В настоящее время традиционные методы выявления мошенничества не всегда могут справиться с новыми угрозами. Традиционные подходы опираются на заранее установленные правила: если транзакция выходит за установленные границы, система подает сигнал. Однако, с появлением новых видов атак, такие системы не способны своевременно обнаружить мошеннические действия.

Это создает необходимость в более адаптивных и умных решениях - машинное и глубокое обучение. Эти технологии способны обрабатывать огромные массивы данных, выявляя скрытые закономерности. Что немаловажно, они могут адаптироваться и обучаться в процессе работы, что делает их эффективными в борьбе с постоянно меняющимися мошенническими стратегиями.

В то же время, актуальным является вопрос эффективности методов глубокого обучения при обнаружении мошеннических транзакций. Для сравнения предлагается рассмотреть основные архитектуры - RNN, CNN и автоэнкодеры, которые, согласно теоретическим материалам, хорошо справляются с распознаванием сложных закономерностей.[1] Также важно найти наиболее эффективный подход для выявления мошенничества, учитывая точность и производительность. Нейронные сети, благодаря своей способности моделировать сложные закономерности и взаимосвязи в данных, могут быть использованы для обнаружения мошеннических транзакций. Кроме того, представляется целесообразным рассмотреть возможность использования ансамблевой модели, объединяющей различные методы глубокого обучения, которая ранее не нашла широкого применения, но может быть перспективной в обнаружении мошеннических транзакций или в других областях, использующих искусственный интеллект. Объединение различных моделей глубокого обучения может привести к созданию более точных

и гибких методов обнаружения мошенничества, способных работать с различными типами данных и их временными зависимостями.

Кроме того, важным аспектом проверки эффективности моделей машинного обучения является создание правильного набора данных, что затрудняется из-за конфиденциальности, которую устанавливают финансовые организации. Основной проблемой синтетических наборов данных является возможное наличие неправильных связей и несбалансированности, что может исказить результаты эксперимента и привести к неправильным выводам об эффективности различных архитектур глубокого обучения.

Тем не менее, в открытых источниках имеются сформированные синтетические наборы финансовых данных, которые, после устранения классового дисбаланса, допустимо использовать для проведения предварительной проверки моделей машинного обучения. Кроме того, на платформе Kaggle доступны реальные анонимизированные наборы данных о транзакциях одной из африканских стран, которые также могут внести вклад в анализ эффективности моделей глубокого обучения. [2]

В данном исследовании предлагается объединить три модели (CNN, RNN, автоэнкодер) по следующей схеме:

- все три модели используют одни и те же входные данные;
- каждая модель обрабатывает информацию по-своему: CNN работает с пространственными признаками, RNN - с последовательностями, а автоэнкодер - сжимает и восстанавливает данные;
- результаты всех трех моделей объединяются на стекинг-слое, который формирует окончательное предсказание.

Такая архитектура позволяет объединить лучшие качества каждой модели и достичь более высокой точности.

Ниже приведены результаты сравнения рассматриваемых моделей глубокого обучения.

Таблица 1 – Результаты сравнения моделей глубокого обучения

№	Наименование модели	Точность обнаружения мошеннических транзакций (%)
1	Сверточная нейронная сеть (CNN)	92,89
2	Рекуррентная нейронная сеть (RNN)	87,5
3	Автоэнкодер	80,1
4	Ансамблевая модель	98

При сравнении выбранных моделей глубокого обучения, лучший результат показала ансамблевая модель (комбинация CNN, RNN и автоэнкодера) с точностью 98%. При анализе используемого набора данных модель использует сильные стороны всех трех методов:

- эффективность CNN для обработки данных с сильными закономерностями;
- применимость автоэнкодера для задач сжатия данных;
- точность RNN при анализе последовательных признаков.

Несмотря на высокую точность результатов ансамблевой модели, при анализе данного подхода были выявлены сложности в оценке производительности и быстродействия, что необходимо учесть при дальнейшем изучении потенциала метода стекирования нейросетевых моделей.

Литература:

1. Аггарвал Ч. Нейронные сети и глубокое обучение: учебный курс – М.: ООО «Диалектика», 2020. – С. 561–565.
2. Khemani, B., Patil, S., Kotecha, K., & Tanwar, S. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. // Journal of Big Data. – 2024. – Vol. 11, N. 1. – P. 18-21.

ИССЛЕДОВАНИЕ МЕХАНИЗМОВ ВНИМАНИЯ В КОМПЬЮТЕРНОМ ЗРЕНИИ В ЗАДАЧАХ КЛАССИФИКАЦИИ

Я.Ю. Собещиков

**Новосибирский государственный технический университет,
г. Новосибирск, sobeshnikov.2020@stud.nstu.ru**

Научный руководитель: Попов А.А., д.т.н., профессор

В работе исследуется гипотеза о том, что нейронные сети с механизмом внимания обеспечивают более высокую точность классификации в задачах компьютерного зрения по сравнению со сверточными нейронными сетями. Рассматриваются различные механизмы внимания и их сравнение с базовыми моделями. Реализованы современные механизмы внимания и построены классификаторы для оценки их точности.

The paper investigates the hypothesis that neural networks with an attention mechanism provide higher classification accuracy in computer vision tasks compared to convolutional neural networks. Various mechanisms of attention and their comparison with basic models are considered. Modern attention mechanisms have been implemented and classifiers have been built to assess their accuracy.

Сверточные нейронные сети, использующиеся сейчас для задач компьютерного зрения, не всегда дают желаемого уровня точности. Одним из вариантов решения данной проблемы – научить сеть фокусироваться на важных участках изображения, игнорируя другие. Это можно назвать механизмом внимания.

Идея механизма внимания была взята из наблюдений за зрительной системой человека. Механизм можно рассматривать как процесс отбора наиболее важных признаков путем адаптивного взвешивания.

В рамках работы были изучены и реализованы механизмы внимания, представленные в таблице 1.

Таблица 1 – Виды внимания и архитектуры

Вид внимания	Архитектуры
Канальное	<i>SE</i> [1], <i>GCT</i> [2], <i>ECA</i> [3], <i>FCA</i> [4]
Пространственное	<i>GE</i> [5]
Канально-пространственное	<i>SCN</i> [6], <i>TripletAttention</i> [7]
Блочное	<i>DynamicConv</i> [8]

В рамках текущей работы были реализованы данные механизмы внимания в виде слоев и построены нейронные сети с ними.

1. Данные

Для исследования точности были использованы следующие открытые датасеты:

- CIFAR10 – набор данных, состоящий из 10 классов животных и средств передвижения. Всего в датасете 60000 экземпляров размером 32x32 пикселя.
- CIFAR100 – набор данных, в котором представлены 100 классов наиболее распространенных объектов окружающего мира. В датасете также 60000 экземпляров размером 32x32 пикселя
- MNIST – набор данных, в котором представлены рукописные цифры от 0 до 9. В датасете 70000 экземпляров размером 28x28 пикселей.

2. Архитектуры нейронных сетей

В качестве базовых моделей для датасетов CIFAR10 и CIFAR100 была представлена следующая архитектура: три одинаковых блока на рис.1 и в конце стандартный блок, состоящий из полносвязных слоев и слоев нормализации.

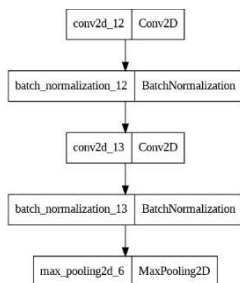


Рис. 10 - Блок архитектуры сети для датасетов CIFAR10 и CIFAR100

В качестве базовой модели для датасета MNIST была выбрана архитектура, изображенная на рисунке 2.

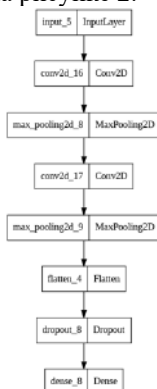


Рис. 11 - Архитектура сети на основе датасета MNIST

При реализации архитектур с механизмами внимания SE, GE после каждого слоя Conv2d были добавлены соответствующие слои.

При реализации архитектур с механизмами внимания FCA, GCT, ECA, TripletAttention перед каждым слоем MaxPooling2D был добавлен соответствующий слой.

При реализации архитектур с механизмом внимания SCN данный слой был вставлен между первым слоем Conv2d и последним слоем MaxPooling2D в каждом блоке.

При реализации архитектур с механизмом внимания DynamicConv все слои Conv2d были заменены на слои внимания.

3. Обучение нейронных сетей

При обучении использовался оптимизатор ADAM для обновления

весов. Для датасетов CIFAR10, CIFAR100 была использована архитектура `sparse_categorical_crossentropy`, для датасета MNIST - `categorical_crossentropy`.

4. Результаты

В результате обучения были получены модели со следующими метриками (см. Таблица 2 – Таблица 4). В результате обучения моделей можно заметить, что использование механизмов не дает большого прироста точности. Можно предположить, что в задаче классификации данные механизмы слабо помогают нейронным сетям обобщать закономерности. Механизмы внимания, представленные в работе, использовались разработчиками для дообучения нейронных сетей, у которых уже была высокая способность к обобщению. При реализации нейронных сетей с нуля, маловероятно, что данные механизмы значительно улучшат ключевые метрики в задаче классификации.

Таблица 2 - Результаты обучения моделей на датасете CIFAR10

Модель	Эпохи	Потери на проверочной выборке	Точность на проверочной выборке, %	Потери на тестовой выборке	Точность на тестовой выборке, %
Базовая	58	0.0056	99.82	0.9085	84.46
SE	54	0.0061	99.8	0.9103	84.52
FCA	54	0.0041	99.87	0.9047	84.09
GCT	37	0.0098	99.7	0.8124	83.66
ECA	58	0.0035	99.92	0.8779	83.31
GE	58	0.0076	99.76	0.9086	85.82
SCN	83	0.0334	98.99	1.6341	81.04
Triplet Attention	52	0.0059	99.83	0.8271	83.67
Dynamic Conv	99	0.1787	99.22	1.2281	81.90

Таблица 3 - Результаты обучения моделей на датасете CIFAR100

Модель	Эпохи	Потери на проверочной выборке	Точность на проверочной выборке, %	Потери на тестовой выборке	Точность на тестовой выборке, %
Базовая	58	0.0084	99.79	3.0554	52.72
SE	70	0.0042	99.88	3.4235	54.25
FCA	78	0.0034	99.91	3.3081	54.44
GCT	56	0.0072	99.85	2.9390	53.97
ECA	58	0.0105	99.82	3.1813	48.80
GE	59	0.0042	99.89	3.2739	53.90
SCN	46	0.0945	97.65	3.7024	47.20
Triplet Attention	56	0.0045	99.89	2.9193	54.76
Dynamic Conv	114	0.2785	99.51	3.6990	51.43

Таблица 4 - Результаты обучения моделей на датасете MNIST

Модель	Эпохи	Потери на проверочной выборке	Точность на проверочной выборке, %	Потери на тестовой выборке	Точность на тестовой выборке, %
Базовая	41	0.0028	99.94	0.0261	99.29
SE	50	0.0021	99.95	0.0278	99.32
FCA	48	0.0021	99.94	0.2024	99.36
GCT	40	0.0029	99.92	0.0259	99.34
ECA	44	0.0016	99.97	0.0282	99.28
GE	22	0.0271	99.33	0.0271	99.33
SCN	27	0.0036	99.89	0.0228	99.35
Triplet Attention	46	0.0017	99.96	0.0229	99.35
Dynamic Conv	108	0.0302	99.94	0.0302	99.94

Литература:

1. Jie Hu, Li Shen, Samuel Albanie, Gang Sun, Enhua Wu. Squeeze-and-Excitation Networks [Electronic resource] // CVPR 2018, 18-22 June 2018: Proceedings. – Salt Lake City, 2018. – P. 7132–7141. – URL: <https://arxiv.org/abs/1709.01507> (accessed: 15.10.2023).
2. Z. Yang, L. Zhu, Y. Wu, Y. Yang. Gated Channel Transformation for Visual Recognition [Electronic resource] // CVPR 2020, 14-19 June 2020:

Proceedings. – Seattle, 2020. – P. 11794–11803. – URL: <https://arxiv.org/pdf/1909.11519>(accessed: 25.10.2023).

3. *Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo, Qinghua Hu*. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks [Electronic resource] // CVPR 2020, 14-19 June 2020: Proceedings. – Seattle, 2020. – P. 11534–11542. – URL: <https://arxiv.org/abs/1910.03151v2> (accessed: 04.11.2023).

4. *Z. Qin, P. Zhang, F. Wu, X. Li*. FcaNet: Frequency Channel Attention Networks [Electronic resource] // ICCV 2021, 11-17 October 2021: Proceedings. – Montreal, 2021. – P.783–792. – URL: <https://arxiv.org/pdf/2012.11879> (accessed: 21.11.2023).

5. *Jie Hu, Li Shen, Samuel Albanie, Gang Sun, Andrea Vedaldi*. Gather-Excite: Exploiting Feature Context in Convolutional Neural Networks [Electronic resource] // NeurIPS 2018, 2-8 December 2018: Proceedings. – Montreal, 2018. – P.1–14. – URL: <https://arxiv.org/abs/1810.12348> (accessed: 15.12.2023).

6. *Jiang-Jiang Liu, Qibin Hou, Ming-Ming Cheng, Changhu Wang, Jiashi Feng*. Improving Convolutional Networks with Self-Calibrated Convolutions [Electronic resource] // CVPR 2020, 14-19 June 2020: Proceedings. – Seattle, 2020. – P.10096–10105. – URL: <https://mftp.mmcheng.net/Papers/20cvprSCNet.pdf> (accessed: 15.01.2024).

7. *Diganta Misra, Trikey Nalamada, Ajay Uppili Arasanipalai, Qibin Hou*. Rotate to Attend: Convolutional Triplet Attention Module [Electronic resource] // WACV 2021, 5-9 January 2021: Proceedings. – Waikoloa, 2021. – P.3139–3148. – URL: <https://arxiv.org/pdf/2010.03045> (accessed: 23.01.2024).

8. *Yinpeng Chen, Xiyang Dai, Mengchen Liu, Dongdong Chen, Lu Yuan, Zicheng Liu*. Dynamic Convolution: Attention over Convolution Kernels [Electronic resource] // CVPR 2020, 14-19 June 2020: Proceedings. – Seattle, 2020. – P.11030–11039. – URL: <https://arxiv.org/abs/1912.03458> (accessed:01.02.2024).

РАЗРАБОТКА МОБИЛЬНОГО ПРИЛОЖЕНИЯ С ПРИМЕНЕНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ ПОТЕНЦИАЛЬНО ОПАСНЫХ КОМПОНЕНТОВ В СОСТАВЕ КОСМЕТИЧЕСКИХ СРЕДСТВАХ ПО СНИМКУ ЭТИКЕТКИ ТОВАРА

Ю.Д. Филипенко

**Новосибирский государственный технический университет,
г. Новосибирск, filipenko.yulia2018@yandex.ru
Научный руководитель: Попов А. А., д.т.н., профессор**

В работе рассматривается проблема определения потенциально опасных компонентов в составе косметических средств. Для решения этой задачи предлагается разработать мобильное приложение с применением методов машинного обучения, которое позволит анализировать состав косметики по снимку упаковки. Исследование включает в себя анализ существующих подходов к определению безопасности косметических средств, разработку алгоритмов машинного обучения для автоматизированного выявления потенциально опасных компонентов, распознавания и классификации ингредиентов, а также создание пользовательского интерфейса мобильного приложения. Результаты работы могут быть полезны для потребителей, производителей косметики и регулирующих органов, так как позволят повысить осведомлённость о составе косметических продуктов и способствовать более ответственному выбору и производству косметики.

The paper considers the problem of determining potentially dangerous components in the composition of cosmetics. To solve this problem, it is proposed to develop a mobile application using machine learning methods that will allow you to analyze the composition of cosmetics based on a package snapshot. The research includes an analysis of existing approaches to determining the safety of cosmetics, the development of machine learning algorithms for automated identification of potentially dangerous components, recognition and classification of ingredients, as well as the creation of a user interface for a mobile application. The results of the work can be useful for consumers, cosmetics manufacturers and regulatory authorities, as they will increase awareness about the composition of cosmetic products and promote more responsible selection and production of cosmetics.

В современном мире косметика стала неотъемлемой частью жизни многих людей. Однако не все знают о том, какие компоненты входят в состав косметических средств и как они могут влиять на здоровье человека. В связи с этим возникает необходимость в разработке инструментов, которые помогут потребителям делать осознанный выбор при покупке косметики. Одним из таких инструментов может стать мобильное приложение, которое будет определять потенциально опасные компоненты в составе косметических средств по снимку.

Приложение для распознавания опасных компонентов в косметике по фото представляет собой инновационный инструмент, который может значительно повысить безопасность потребителей, улучшить качество косметических продуктов и предоставить множество полезных функций различным заинтересованным сторонам. В этой статье рассматриваются основные области применения разрабатываемого приложения и анализируются, как оно может быть использовано в разных контекстах для достижения максимальной пользы.

Анализ существующих приложений показывает, что каждое из них имеет свои сильные и слабые стороны. Основные функции включают фотографирование этикеток, использование OCR для распознавания текста и анализ состава с предоставлением информации о безопасности ингредиентов [1]. Однако приложение может либо не поддерживать русский язык, либо распознавать продукт только по штрих-коду. Таким образом создание русифицированного приложения для распознавания опасных компонентов имеет большие перспективы использования на всеобщем рынке.

В контексте разработки данного приложения, предназначенного для распознавания опасных компонентов в косметике по фото, библиотека Tesseract.OCR для C# играет ключевую роль [2]. С ее помощью можно извлекать текстовую информацию с изображений этикеток косметических продуктов.

После завершения этапа обучения системы распознавания текста можно переходить к процессу выявления компонентов в составе косметических продуктов [3]. Результаты распознавания текста предоставляют возможность автоматизированного анализа этикеток, а затем выделения и идентификации вредных компонентов.

Таблица 1 – Опасные косметические компоненты

Название компонента	Обозначение в составе	Влияние на здоровье
Фталаты	DEP, DBP, DMP	Эндокринные расстройства, раздражение кожи, возможные канцерогенные эффекты
Парабены	Methylparaben, Ethylparaben, Propylparaben	Возможные нарушения гормонального баланса, раздражение, аллергические реакции.
Сульфаты	SLS, SLES	Раздражение кожи, сухость, потенциальные канцерогенные эффекты.

Таблица 1 – Продолжение

Триклозан	Triclosan	Эндокринные нарушения, воздействие на микробиому кожи.
Формальдегид	Formaldehyde, Methanal, Formalin	Раздражение, аллергии, потенциально канцерогенное действие.
Октокрилен	Octocrylene	Возможное нарушение гормонального баланса, раздражение, аллергии.
Минеральные масла	Mineral Oil, Paraffinum Liquidum, Petrolatum	Засорение пор, ухудшение дыхания кожи, раздражение.
Этаноламин	DEA, MEA, TEA	Раздражение, аллергии, связано с образованием канцерогенных соединений.
Гидроксианизол	Butylated Hydroxyanisole	Подозрение на канцерогенность, раздражение, аллергии
Оксибензон	Oxybenzone, Benzophenone-3	Подозрение на эндокринные нарушения, ассоциировано с риском развития рака.
Акриламид	Acrylamide	Подозрение на канцерогенность, особенно в продуктах для ухода за волосами.

В данной таблице представлены лишь некоторые опасные компоненты, однако для более детального выявления вредных составляющих, необходимо интегрировать базу данных, включающую большое количество опасных компонентов, их обозначение и влияние на здоровье [4].

После выявления опасных компонентов можно приступить к интеграции собранной базы данных в модель по распознаванию текста. Этот важный этап позволяет улучшить функциональность модели, обеспечивая более глубокий анализ состава косметических продуктов [5]. Интегрированная база данных будет содержать информацию о вредных и запрещенных компонентах, их наименованиях и влиянии на здоровье (Рисунок 1).

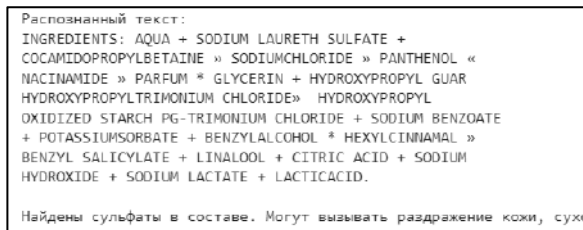


Рис. 1 – Выявление компонентов

Таким образом была получена модель, способная не только распознать текст на изображении этикетки, но и, проанализировав его, выдать список опасных или нежелательных компонентов.

В заключении статьи можно отметить, что использование методов машинного обучения для распознавания текста на упаковках косметических продуктов предоставляет перспективные возможности в области обеспечения безопасности и качества товаров. Однако, проведенный анализ показывает, что эффективность модели напрямую зависит от качества обучающих данных и их соответствия реальным условиям применения.

Выявлены некоторые трудности в распознавании текста на упаковках, такие как форма тары, что искривляет текст и делает его сложным для распознавания. Для более точного и надежного распознавания необходимо продолжить улучшение обучающих данных.

Несмотря на выявленные недостатки, результаты исследования подтверждают потенциал применения методов машинного обучения в области контроля качества косметических продуктов. Таким образом, модель становится более функциональной и эффективной инструментом для пользователей, желающих принимать осознанные решения относительно выбора косметики и ухода за кожей.

Литература:

1. Корниенко С.И., Черепанов Ф.М., Ясницкий Л.Н. Распознавание текстов рукописных и старопечатных книг на основе нейросетевых технологий – М., 2008г.
2. Полное руководство по OCR с Tesseract, OpenCV и Python [электронный ресурс]. - <https://zephyrnet.com/ru/полное-руководство-по-ocr-с-тессерактом-opencv-и-python/> (дата обращения 28.10.2024).
3. Пучкова Т. Энциклопедия косметических ингредиентов. Школа косметических химиков – М., 2015. – 336 с.

4. Башура А.Г., Половоко Н.П. Технология косметических и парфюмерных средств - во. НФАУ «Золотые страницы», 2002 г., 272 с.

5. Ян Эрик Солем. Программирование компьютерного зрения на языке Python – 2012. – 314 с.

ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ ЗАБОЛЕВАНИЙ КОЖИ ПО ДЕРМАТОСКОПИЧЕСКИМ ИЗОБРАЖЕНИЯМ

А.П. Хайдуков

**Новосибирский государственный технический университет,
г. Новосибирск, hajdukov.2020@stud.nstu.ru**

Научный руководитель: Попов А.А., д.т.н., профессор

В данной работе рассматривается актуальная проблема распознавания заболеваний кожи с помощью компьютерного зрения и искусственного интеллекта. Проводится сравнение и исследование эффективности моделей, основанных на архитектурах EfficientNet и EfficientNetV2 на наборе дерматоскопических изображений.

This paper addresses the current problem of skin disease recognition using computer vision and artificial intelligence. A comparison and study of the effectiveness of models based on the EfficientNet and EfficientNetV2 architectures is conducted on a set of dermoscopic images.

Распознавание дерматоскопических изображений является важной задачей в диагностике кожных заболеваний. В последнее время нейронные сети показали высокую эффективность в решении этой задачи. В данной статье исследуется использование архитектур EfficientNet и EfficientNetV2 для распознавания дерматоскопических изображений на основе набора данных ISIC-2019. Проведена серия экспериментов для оценки влияния дисбаланса данных на точность классификации и исследованы различные методы, такие как отбрасывание изображений из мажоритарного класса и аугментация данных, для улучшения точности классификации.

Для создания модели были выбраны архитектуры EfficientNet [1] и EfficientNetV2 [2].

Архитектура EfficientNet разработана с использованием метода автоматического проектирования нейронных сетей, названного Neural Architecture Search (NAS) [3]. Авторы использовали метод, названный составным коэффициентом масштабирования (рисунок 1), для

масштабирования базовой модели EfficientNet-B0 и разработали несколько вариантов EfficientNet, от EfficientNet-B1 до EfficientNet-B7.

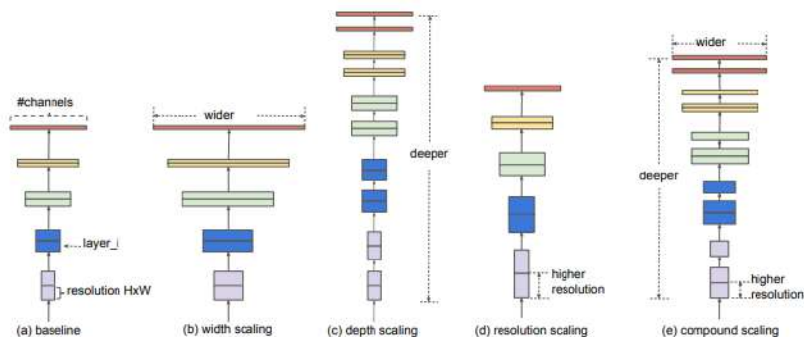


Рис. 1 – Иллюстрация методов масштабирования [1]

Суть составного масштабирования заключается в сбалансированном масштабировании сети. При масштабировании сети в 2^N раз, ее глубина увеличивается в α^N раз, ширина в β^N раз и размер входных изображений в γ^N раз, где α , β и γ это коэффициенты определенные с помощью большого количества тестов с малым шагом на базовой модели малого размера.

EfficientNetV2 является улучшенной версией предыдущей архитектуры, в которой исправлены проблемы с неоптимальным распределением вычислительных ресурсов и узкие места, замедляющие обучение.

Для обучения модели использовался набор изображений кожных заболеваний ISIC-2019. Этот набор содержит 7 классов дерматоскопических изображений и состоит из более чем 21 тысячу изображений. Датасет не сбалансирован, объем наибольшего класса превышает объем наименьшего в 54 раза.

В ходе исследования для создания моделей нейронных сетей использовались такие архитектуры как EfficientNetB0, EfficientNetB3 и EfficientNetV2-S, которые обеспечивают хорошее соотношение между качеством классификации и потребляемыми вычислительными мощностями, измеряемыми в FLOPS (количество операций с плавающей запятой в секунду).

В ходе исследования было проведено 4 серии экспериментов.

1. Обучение моделей на оригинальном, несбалансированном датасете;

2. Обучение моделей на датасете, из которого отброшены 70% изображений из мажоритарного класса;

3. Обучение моделей на датасете, из которого отброшены 55% изображений из мажоритарного класса, и остальные классы увеличены в 1,5 раза;

4. Обучение со взвешиванием на датасете, из которого отброшены 55% изображений из мажоритарного класса, и остальные классы увеличены в 1,5 раза.

Каждая серия экспериментов была направлена на улучшение точности моделей за счет снижения дисбаланса датасета и его влияния на обучение модели. Все изображения из набора данных сжимались до размера 364 на 384 пикселя. Для третьей и четвертой серии экспериментов аугментация [4] производилась путем случайного поворота на угол до 45 градусов, отражение по вертикали и горизонтали и другие преобразования изображения.

В качестве оптимизатора использовался алгоритм AdamW [5].

Низкая точность классификации в случае с датасетом ISIC-2019 объясняется несбалансированностью, что приводит к тому, что нейросеть хорошо угадывает изображения мажоритарного класса, и сильно хуже остальные классы. При отбрасывании 70% мажоритарного класса дисбаланс стал меньше, но и объем датасета уменьшился на 35%, что также негативно влияет на узнавание практически всех классов изображений. При отбрасывании 55% мажоритарного класса и дополнении датасета аугментированными изображениями, его объем остался прежним, при этом дисбаланс значительно снизился, но это помогло добиться лишь небольшого улучшения результата. Предположительно негативным фактором в данном случае оказалось то, что датасет был дополнен похожими изображениями и нейросеть переобучилась. Взвешивание классов позволило в среднем увеличить точность классификации и уменьшить разницу в качестве распознавания различных классов, что видно по значениям метрик для отдельных классов в таблице 2.

Таблица 2 – Метрики обученных моделей

Датасет	Архитектура	Метрики		
		accuracy	macro avg	weighted avg
Оригинальный датасет	EfficientNetB0	0,77	0,72	0,77
	EfficientNetB3	0,8	0,75	0,8
	EfficientNetV2-S	0,81	0,76	0,81
Отбрасывание 70% мажоритарного класса	EfficientNetB0	0,78	0,73	0,78
	EfficientNetB3	0,8	0,76	0,8
	EfficientNetV2-S	0,81	0,77	0,81
Аугментация	EfficientNetB0	0,8	0,76	0,8
	EfficientNetB3	0,81	0,77	0,81
	EfficientNetV2-S	0,83	0,81	0,83
Взвешивание	EfficientNetB0	0,82	0,81	0,81
	EfficientNetB3	0,83	0,83	0,83
	EfficientNetV2-S	0,86	0,86	0,86

Таблица 2 – Метрики классов взвешенного обучения

Класс	precision			recall			f1-score		
	B0	B3	V2-S	B0	B3	V2-S	B0	B3	V2-S
Актин. кератоз	0,61	0,63	0,79	0,72	0,72	0,71	0,66	0,67	0,75
Базалома	0,86	0,82	0,87	0,82	0,88	0,95	0,84	0,85	0,91
Доброкач. Кератоз	0,68	0,71	0,79	0,77	0,74	0,84	0,72	0,72	0,81
Дермато-фибома	0,39	0,53	1,00	0,64	0,73	0,82	0,48	0,62	0,9
Родинка	0,9	0,92	0,85	0,85	0,85	0,86	0,87	0,88	0,86
Меланома	0,7	0,74	0,88	0,76	0,8	0,8	0,73	0,77	0,84
Сосудистые образования	0,85	0,9	0,92	0,71	0,79	0,96	0,77	0,84	0,94

В результате исследования было обнаружено, что несбалансированность данных является серьезной проблемой для распознавания кожных заболеваний по дерматоскопическим изображениям. Исследование показывает, что использование нейронных сетей, таких как EfficientNet и EfficientNetV2, в сочетании с методами обработки дисбаланса данных может существенно улучшить

точность распознавания изображений кожи. Результатом проектирования стала модель с архитектурой EfficientNetV2-S, обученная на взвешенном датасете ISIC-2019 с дополнением малых классов, достигшая лучшей точности классификации в 86%.

Литература:

1. Tan, M. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks / M. Tan, Q. Le // Proceedings of Machine Learning Research. – 2019. – № 97. – С. 6105-6114.
2. Tan, M. EfficientNetV2: Smaller Models and Faster Training / M. Tan, Q. Le // Proceedings of Machine Learning Research. – 2021. – № 139. – С. 10096-10106.
3. Neural architecture search with reinforcement learning / B. Zoph, Q. V. Le // Arxiv.org: электронный журнал. – URL: <https://doi.org/10.48550/arXiv.1611.01578> (accessed 28.10.2024).
4. Data Augmentation as Feature Manipulation / R. Shen, S. Bubeck, S. Gunasekar // Arxiv.org: электронный журнал. – URL: <https://doi.org/10.48550/arXiv.2203.01572> (accessed 27.10.2024).
5. Understanding AdamW through Proximal Methods and Scale-Freeness / Z. Zhuang, M. Liu, A. Cutkosky [и др.] // Arxiv.org: электронный журнал. – URL: <https://doi.org/10.48550/arXiv.2202.00089> (accessed 29.10.2024).

АНАЛИЗ ИСПОЛЬЗОВАНИЯ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА СТУДЕНТАМИ ДЛЯ РЕШЕНИЯ ЗАДАНИЙ ПО ПРОГРАММИРОВАНИЮ

Е.Р. Чеглов

Новосибирский государственный университет

экономики и управления «НИНХ»,

г. Новосибирск, e.for.work@yandex.ru

Научный руководитель: Пестунов А.И., к.ф.-м.н., доцент

В работе рассматривается влияние генеративного искусственного интеллекта на учебный процесс по программированию. В ходе исследования, включающего анкетирование 61 студента, выявлены паттерны использования ИИ, а также его преимущества и недостатки в учебной деятельности. Установлено, что ИИ помогает студентам ускорить процесс освоения материала, улучшая понимание сложных алгоритмов и повышая уверенность в собственных силах. Однако выявлены и риски: излишняя зависимость от ИИ может снижать уровень критического мышления и мотивацию к

самостоятельному решению задач. Полученные результаты подчеркивают необходимость разработки методик, направленных на эффективное использование ИИ без ущерба для самостоятельности студентов.

The paper considers the influence of generative artificial intelligence on the educational process of programming. In the course of the study, including questionnaire survey of 61 students, the patterns of AI use, as well as its advantages and disadvantages in learning activities were identified. It was found that AI helps students to speed up the process of mastering the material, improving their understanding of complex algorithms and increasing their self-confidence. However, risks were also identified: over-reliance on AI can reduce critical thinking and motivation for independent problem solving. These results emphasise the need to develop methods aimed at the effective use of AI without compromising student autonomy.

С развитием генеративного искусственного интеллекта (ИИ) студенты получили доступ к инструментам, позволяющим автоматизировать выполнение заданий по различным дисциплинам, в том числе по программированию. Генеративный ИИ позволяет создавать решения, генерируя корректный и понятный программный код, что заметно снижает временные и ресурсные затраты на выполнение заданий. В то же время использование генеративного ИИ в образовательном процессе вызывает активные дискуссии среди преподавателей и исследователей. С одной стороны, ИИ может служить личным ассистентом для студентов, объясняя алгоритмы или процессы разработки сложных программных конструкций, а также аккумулируя информацию в удобной форме [1]. С другой — существует опасение, что использование генеративного ИИ может существенно снизить мотивацию к самостоятельному поиску решений и помешать развитию критического и креативного мышления, что в перспективе может негативно сказаться на уровне знаний студентов [2].

В связи с этим возникает ряд вопросов: является ли использование генеративного ИИ угрозой для образовательного процесса? Может ли он способствовать более глубокому пониманию сложных тем, освободив время для анализа и творчества, или же, напротив, лишает студентов необходимого практического опыта? Цель данного исследования — провести анализ опыта студентов в использовании генеративного ИИ для решения учебных заданий по программированию. Исследование направлено на выявление характерных особенностей и паттернов применения генеративного ИИ среди студентов различных категорий, что позволит лучше понять, как использование ИИ влияет на образовательный процесс в контексте обучения программированию.

В данном исследовании анализируются мнения и опыт студентов, применяющих различные модели генеративного ИИ для выполнения задач по программированию. В рамках исследования разработана анкета для сбора количественных данных, включающая также открытые вопросы с целью получения качественной информации.

Анкетирование проводилось среди студентов, изучающих программирование в рамках дисциплин: «Основы алгоритмизации и программирования» (среднее профессиональное образование) и «Программирование», «Языки программирования», «Программирование дискретных структур» (бакалавриат). Участие в исследовании носило добровольный и анонимный характер. Всего в опросе принял участие 61 студент, из которых 29 обучались на программах среднего профессионального образования, а 32 — на программах бакалавриата.

Согласно результатам анкетирования, 34 студента (55,7%) указали, что обладают лишь базовыми знаниями в программировании, а 11 студентов (18%) отметили, что изучают программирование впервые. Ещё 11 респондентов (18%) оценили свои навыки как уверенные на одном языке программирования, и только 5 студентов (8,2%) сообщили, что способны писать код на нескольких языках программирования.

При самооценке владения алгоритмами 29 студентов (47,5%) заявили об отсутствии знаний в данной области, тогда как 26 студентов (42,6%) указали, что знакомы с некоторыми базовыми алгоритмами, такими как полный перебор или сортировка пузырьком. Лишь 6 студентов (9,8%) отметили владение более сложными алгоритмами, например, бинарным поиском, быстрой сортировкой или алгоритмом Евклида.

Анализ ответов респондентов позволяет выделить две группы студентов в зависимости от уровня их владения программированием и алгоритмами. Среди студентов, которые только начали изучать программирование или обладают минимальными базовыми знаниями, наибольшую популярность использование генеративного ИИ получило из-за трудностей, возникающих после нескольких неудачных попыток самостоятельного выполнения задания (67%), а также необходимости исправления ошибок в собственном коде (65%). Напротив, студенты, обладающие более уверенными навыками программирования на одном или нескольких языках, чаще обращались к ИИ из-за непонимания условий или смысла задачи (44%) и недостатка мотивации для самостоятельного написания решения (38%). Также для обеих групп отмечена такая причина, как стремление получить оптимизированное решение (13,3% от общего числа респондентов).

Анализ частоты использования генеративного ИИ показал, что 73,8% опрошенных применяют его 1–3 раза в неделю, 16,4% обращаются к ИИ ежедневно, и 9,8% используют его 4–5 раз в неделю.

Студенты отметили, что использование генеративного ИИ при выполнении заданий по программированию даёт ряд значительных преимуществ. Наиболее часто упоминаемыми среди них стали экономия времени, которую отметили 82,2% опрошенных, и помощь в освоении новых и сложных, с точки зрения самих студентов, алгоритмов, которую отметили 74,2% респондентов. Среди дополнительных преимуществ применения генеративного ИИ студенты отметили его роль в ускорении и улучшении процесса освоения программирования, более доступные и понятные объяснения и решения по сравнению с традиционными объяснениями преподавателя, а также повышение уверенности в собственных силах и рост мотивации к обучению.

Некоторые из студентов (2,9%) не видят недостатков в использовании генеративного ИИ при выполнении заданий по программированию. Однако большинство респондентов отметили ряд существенных негативных аспектов. Студенты начального уровня, имеющие минимальные базовые знания, наиболее часто указывали на неточность сгенерированного кода (72%) и риск привыкания к использованию ИИ (44%), что, по их мнению, ослабляет способности к алгоритмическому мышлению (35%). Студенты с более высоким уровнем подготовки также отмечали частую неправильность генерируемых решений (88%) и проблему неверного интерпретирования ИИ их запросов, что зачастую вводит в заблуждение (50%). К другим недостаткам респонденты отнесли отсутствие творческого подхода и вариативности в решении задач.

Результаты опроса показывают, что большинство студентов (61%) указывают на необходимость нескольких уточняющих запросов для получения корректного решения от генеративного ИИ, хотя 34% респондентов отмечают успешное решение задачи с первого запроса. Эти данные могут свидетельствовать о том, что многие студенты ещё не освоили навыки точной формулировки запросов к ИИ и склонны просто копировать условие задачи, отправляя его на обработку без адаптации.

Кроме того, опрос показал, что студенты редко представляют сгенерированные решения как свои собственные (лишь 11%). Чаще они используют ИИ для улучшения собственных решений: 43% респондентов интегрируют часть сгенерированного кода в свою работу, а 46% вносят изменения в предложенный код перед отправкой на проверку.

Полученные результаты позволяют оценить степень возможности интеграции генеративного ИИ в образовательный процесс. С одной стороны, ИИ может выступать как инструмент для индивидуальной поддержки студентов, учитывая их уровень подготовки и потребности. С другой стороны, он может негативно повлиять на развитие самостоятельного мышления. Полученные результаты подчеркивают необходимость дальнейших исследований в области интеграции ИИ в образовательный процесс, особенно с акцентом на создание методик, которые помогут студентам развивать навыки точной постановки задач и эффективного использования ИИ как учебного инструмента, не снижая при этом навыки критического мышления у студентов и мотивацию к их самостоятельной работе.

Литература:

1. *Ханиев Р.М.* Использование инновационных технологий в процессе обучения в высшей школе // Проблемы и перспективы развития социально-экономических и гуманитарных наук: педагогика, психология, экономика, юриспруденция: сб. науч. ст. – Саратов: Изд-во Центр вынужденных переселенцев «Саратовский источник», 2024. – С. 119–124.
2. *Валиахметова Н.Р., Ахмадуллина Р.М., Ярмакеев И.Э.* Возможности и риски применения нейростей в образовании // Филология и культура. – 2024. – №2. – С. 260–271.

ИНТЕГРАЦИЯ ГЕЙМИФИКАЦИИ И ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ПОВЫШЕНИЯ ВОВЛЕЧЕННОСТИ УЧЕНИКОВ: ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ НА БАЗЕ ОБРАЗОВАТЕЛЬНОЙ ОНЛАЙН-ПЛАТФОРМЫ

И.М. Шайхисламов

Новосибирский государственный университет

экономики и управления НИНХ»,

г. Новосибирск, ildar.inkin@gmail.com

Научный руководитель: Пестунов А.И., к.ф.-м.н., доцент

В статье исследуется влияние интеграции геймификации и технологий искусственного интеллекта на вовлеченность учеников в онлайн-обучении. Эксперимент с 1000 участниками показал, что использование игровых элементов и адаптивного обучения значительно увеличивает успеваемость и удовлетворенность обучающихся. Результаты продемонстрировали

статистически значимые улучшения в ключевых показателях, подчеркивая потенциал синергетического эффекта от сочетания геймификации и ИИ в образовательных технологиях.

This article explores the impact of integrating gamification and artificial intelligence technologies on learner engagement in online learning. An experiment with 1000 participants showed that the use of game elements and adaptive learning significantly increased learner performance and satisfaction. The results showed statistically significant improvements in key indicators, highlighting the potential for synergies between gamification and AI in educational technology.

В эпоху цифровой трансформации образования онлайн-обучение становится все более распространенным. Однако одной из ключевых проблем, с которой сталкиваются образовательные платформы, является низкий уровень вовлеченности обучающихся. Это приводит к снижению эффективности обучения и высоким показателям отсева.

Цель данного исследования – оценить влияние интеграции элементов геймификации и технологий искусственного интеллекта на вовлеченность учеников в процесс онлайн-обучения. Основная гипотеза заключается в том, что комбинация этих подходов может значительно повысить уровень вовлеченности учащихся, улучшить их успеваемость и общую удовлетворенность процессом обучения.

Геймификация, определяемая как использование игровых элементов в неигровых контекстах [1], приобрела значительную популярность в образовательной сфере. Исследования показывают, что внедрение игровых механик может повысить мотивацию учеников и улучшить результаты обучения [2].

Элементы геймификации в образовании, такие как системы очков, бейджи, лидерборды, уровни сложности и нарративы способствуют созданию более увлекательной и интерактивной учебной среды.

Применение технологий ИИ в образовании открывает новые возможности для персонализации учебного опыта. Алгоритмы машинного обучения позволяют анализировать паттерны поведения учеников и адаптировать контент под индивидуальные потребности каждого учащегося [3].

Исследования показывают, что ИИ-driven персонализация может значительно повысить эффективность обучения. Системы, основанные на технологиях ИИ, способны предсказывать трудности учеников и предоставлять своевременную поддержку [4].

Несмотря на то, что геймификация и ИИ по отдельности показали свою эффективность в повышении вовлеченности учеников, исследований их совместного применения недостаточно. Данное исследование направлено на заполнение этого пробела, изучая

потенциальный синергетический эффект от интеграции геймификации и технологий ИИ в образовательной платформе.

Для проверки гипотезы проведен контролируемый эксперимент на базе онлайн-платформы с общим числом пользователей 5000 человек. Из них случайным образом были отобраны 1000 учеников для участия в исследовании. Участники были разделены на две группы: экспериментальную и контрольную по 500 учеников в каждой.

Экспериментальная группа использовала версию платформы с интегрированными элементами геймификации и ИИ-driven персонализацией. Контрольная группа использовала стандартную версию платформы. Эксперимент проводился в течение 12 недель.

В экспериментальной версии платформы были реализованы следующие элементы геймификации:

- система очков и уровней;
- бейджи за достижения;
- лидерборды;
- персонажи и нарративы.

ИИ-driven персонализация использовалась для адаптации сложности заданий и предсказания потенциальных трудностей с предоставлением своевременной поддержки.

Для оценки вовлеченности учеников использовались следующие показатели:

- время, проведенное на платформе (T);
- количество выполненных заданий (A);
- регулярность посещений (R);
- успеваемость (P).

На основе этих показателей был разработан комплексный индекс вовлеченности (EI), рассчитываемый по формуле:

$$EI = \frac{0.3 \times T + 0.3 \times A + 0.2 \times R + 0.2 \times P}{4} \quad (1)$$

Все переменные нормализованы к шкале от 0 до 1.

Для реализации адаптивного обучения была использована архитектура многослойного перцептрона (MLP). Сеть состоит из входного слоя, нескольких скрытых слоев и выходного слоя. Количество нейронов в выходном слое соответствует количеству предметов, изучаемых на платформе.

Для обучения модели использовалась функция потерь категориальной кросс-энтропии. Оптимизатор Adam использовался с начальной скоростью обучения $\eta = 0.001$.

После завершения эксперимента мы получили следующие средние значения индекса вовлеченности:

- экспериментальная группа: $EI_{exp} = 0.76$ ($SD = 0.12$)

- контрольная группа: $EI_{control} = 0.58$ ($SD = 0.15$)

Для оценки статистической значимости различий между группами был проведен независимый t-тест:

$$t = \frac{EI_{exp} - EI_{control}}{\sqrt{\frac{s_{exp}^2}{n_{exp}} + \frac{s_{control}^2}{n_{control}}}} \quad (2)$$

где s_{exp} и $s_{control}$ - стандартные отклонения, n_{exp} и $n_{control}$ - размеры выборок.

Результаты t-теста: $t = 4.82$, $d_f = 998$, $p < 0.001$

Полученные результаты указывают на статистически значимое различие в уровне вовлеченности между экспериментальной и контрольной группами.

Все отдельные компоненты индекса вовлеченности показали значительное улучшение в экспериментальной группе, с наибольшим приростом в количестве выполненных заданий.

Таблица 1 – Сравнение показателей вовлеченности

Показатель	Экспериментальная группа	Контрольная группа	Разница
Время (Т)	0.79	0.61	+29.5%
Задания (А)	0.82	0.59	+39.0%
Регулярность (R)	0.74	0.57	+29.8%
Успеваемость (Р)	0.71	0.55	+29.1%

Наблюдаемый синергетический эффект превосходит результаты предыдущих исследований, фокусировавшихся только на геймификации или только на ИИ-персонализации. Статистически значимая разница в индексе вовлеченности между экспериментальной и контрольной группами (0.76 против 0.58) подтверждает исходную гипотезу и эффективность данного подхода.

Результаты данного исследования имеют важное значение для разработчиков образовательных платформ и педагогов, стремящихся повысить эффективность онлайн-обучения. Интеграция геймификации

и технологий ИИ представляется перспективным направлением развития образовательных технологий, способным значительно улучшить опыт онлайн-обучения для учеников.

Литература:

1. *Deterding S., Dixon D., Khaled R. & Nacke L.* From game design elements to gamefulness: defining "gamification" // In Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments, MindTrek 2011, 28-30 sept. 2011: Proceedings. – New York, 2011. – P. 9–15.
2. *Subhash S. & Cudney E.A.* Gamified learning in higher education: A systematic review of the literature // Computers in Human Behavior. – 2018. – Vol. 87. – P. 192-206.
3. *Faridi B., Shaheen S.S.* Online learning platforms and teacher efficacy // International Journal of Humanities and Education Research. – 2024. – Vol. 6(1). – P. 15-24. – doi: 10.33545/26649799.2024.v6.i1a.64.
4. *Прохорова М.П., Шкунова А.А., Гуреева Е.П.* Средства персонализации образовательного процесса в рамках электронных курсов // Проблемы современного педагогического образования. – 2021. – № 71-3. – С. 183-187.

**ИССЛЕДОВАНИЕ СВЯЗИ ЭКОНОМИЧЕСКИХ
И ЭКОЛОГИЧЕСКИХ ПОКАЗАТЕЛЕЙ НСО
С ИСПОЛЬЗОВАНИЕМ ЧАСТНЫХ КОРРЕЛЯЦИОННЫХ СЕТЕЙ**

Г.В. Белозубов

**Новосибирский государственный технический университет,
г. Новосибирск, soacker@yandex.ru**

Научный руководитель: Тимофеева А.Ю., к.э.н., доцент

В данной работе исследуется наличие возможных связей между экономической и экологической обстановкой в Новосибирской области путём построения частных корреляционных сетей по коротким временным рядам. Для однозначной интерпретации результата статистически малозначимые рёбра и изолированные вершины отбрасывались. В результате исследования было определено и обосновано наличие корреляций между рядом экономических и экологических показателей Новосибирской области.

In this paper, the existence of possible links between the economic and environmental situation in the Novosibirsk oblast' is investigated by constructing partial correlation networks over short time series. For an unambiguous interpretation of the result, statistically insignificant and isolated vertices were discarded. As a result of the study, the existence of correlations between a number of economic and environmental indicators of the Novosibirsk oblast' was determined and substantiated.

Степень влияния человека на природу вызывает бурные дискуссии на различных уровнях, от обывателей до учёных-экологов и политиков. Не исключение и Новосибирская область, на территории которой в последние годы наблюдались ухудшение экологической обстановки и аномальные погодные явления. В связи с этим, целью данной работы является исследование взаимосвязей между различными экологическими и экономическими показателями Новосибирской области за период с 2006 по 2022 год.

Экономические показатели включают в себя изменение валового регионального продукта по отраслям, а экологические, в свою очередь, такие показатели, как объёмы загрязняющих веществ, выброшенных в атмосферу и удельный вес проб воды/почвы, не соответствующих нормативам по санитарно-химическим показателям.

Также рассматривались иные показатели: площадь лесных пожаров и количество единиц охотничьих ресурсов (кабан, водоплавающая

птица), однако данные показатели были отброшены из-за отсутствия статистически значимой связи с другими показателями. Исследование проводилось на данных из открытых источников [1, 2] в формате коротких временных рядов с шагом в год.

Для анализа данных были использованы частные корреляционные сети [3]. Преимущество такого подхода заключается в том, что частные корреляционные сети игнорируют вклад ложных корреляций и более точно описывают взаимосвязи между переменными. Работа была выполнена на языке *R* с использованием пакета *qgraph* [4]. Ключевые этапы проведения исследования включали в себя: 1) выбор подмножества признаков со значимыми корреляциями; 2) сбор статистических данных из открытых источников; 3) расчёт матрицы частных корреляций; 4) построение графа частной корреляционной сети; 5) интерпретация полученного графа сети.

В результате, были получены и проинтерпретированы графы частных корреляционных сетей (рис. 1), в связи с чем выявлено наличие значимой взаимосвязи экологической и экономической обстановки в Новосибирской области. Рёбра с уровнем корреляции ниже 0.25 отбрасывались как статистически малозначимые, что было необходимо для упрощения структуры графа для интерпретации результатов.

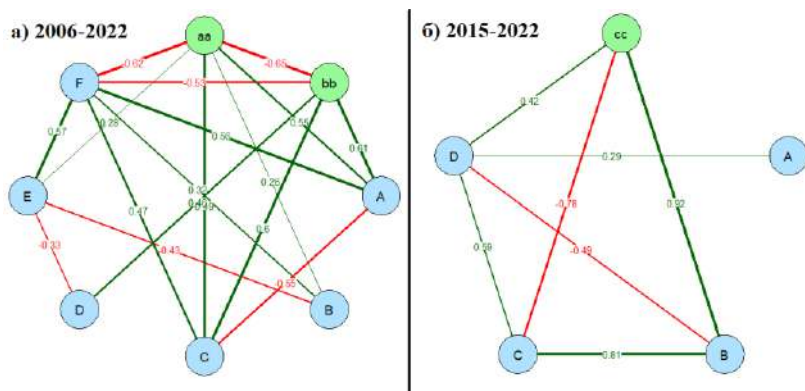


Рис. 1 – Частная корреляционная сеть, 2 экологических показателя и 6 экономических показателей за период 2006-2022 гг. (а); 1 экологический показатель и 4 экономических показателя за период 2015-2022 гг. (б)

Таблица 1 – Легенда для графов частных корреляционных сетей

Экологические показатели	
aa	Объем выбросов загрязняющих веществ в атмосферный воздух, тыс. тонн
bb	Удельный вес проб воды, не соответствующих нормативам по санитарно-химическим показателям, % (Источники централизованного водоснабжения)
cc	Удельный вес проб почвы, не соответствующих нормативам по санитарно-химическим показателям, % (В селитебной зоне)
Экономические показатели	
A	Сельское, лесное хозяйство, охота, рыболовство и рыбоводство
B	Добыча полезных ископаемых
C	Обрабатывающие производства
D	Производство и распределение электроэнергии, газа и воды
E	Строительство
F	Транспортировка и хранение
G	Деятельность гостиниц и предприятий общественного питания

Было отмечено, что вид получаемого графа сильно чувствителен к выбору подмножества сопоставляемых признаков. К примеру, расчёт матрицы частных корреляций для 2-х экологических и 5-и экономических показателей приводил к получению графа, который сильно уступал по возможностям интерпретации другому графу, содержащему 2 экологических и 6 экономических показателей. Также, оптимальным для последующего анализа оказался граф, включающий в себя показатель удельных весов аномальных проб почвы и 4-е экономических показателя.

Из анализа полученных частных корреляционных сетей последовали следующие выводы:

1) темпы роста объёмов выбросов загрязняющих веществ в атмосферу напрямую коррелировали с темпами роста ВРП в отраслях сельского/лесного хозяйства, добычи полезных ископаемых, обрабатывающих производств и строительства;

2) темпы роста удельных весов проб воды, не соответствующих нормативам по санитарно-химическим показателям в источниках централизованного водоснабжения, напрямую коррелировали с темпами роста ВРП в отраслях сельского/лесного хозяйства, обрабатывающих производств и производства/распределения электроэнергии, газа и воды;

3) темпы роста удельных весов проб почвы, не соответствующих нормативам по санитарно-химическим показателям в селитебной зоне,

напрямую коррелировали с темпами роста ВРП в отраслях добычи полезных ископаемых и производства/распределения электроэнергии, газа и воды.

Литература:

1. Территориальный орган Федеральной службы государственной статистики по Новосибирской области: Валовой региональный продукт, Официальная статистика [Электронный ресурс] // ВРП с 1998 года. – URL: <https://54.rosstat.gov.ru/folder/140578> (дата обращения: 05.01.2024).
2. Правительство Новосибирской области [Электронный ресурс] // Государственный доклад о состоянии и об охране окружающей среды Новосибирской области в 2005-2022 г. – URL: <https://www.nso.ru/page/2624> (дата обращения: 05.01.2024).
3. Epskamp S., Fried E. I. A tutorial on regularized partial correlation networks // Psychological methods. – 2018. – Т. 23. – №. 4. – С. 617.
4. Epskamp S. et al. [Электронный ресурс] // Пакет qgraph для R. – URL: <https://cran.r-project.org/web/packages/qgraph/index.html> (дата обращения: 05.01.2024).

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТЕЙ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ИЗМЕРЕНИЯ ПЕРЕХОДНОГО СЛОЯ ПРИ SPS-СПЕКАНИИ

Н.Е. Бортников

**Новосибирский государственный технический университет,
г. Новосибирск, bortnikovnik@gmail.com**

Научный руководитель: Тимофеев В.С., д.т.н., доцент

В данной статье исследуются возможности применения свёрточных нейронных сетей для сегментации переходного слоя по снимкам сплава двух смесей металлов. Реализована нейронная сеть по архитектуре U-Net. На основе показаний двух метрик исследована эффективность применения трёх функций потерь. Результаты сравнены с результатами ручного измерения переходного слоя.

The article explores the potential of using convolutional neural network for segmentation of the transition layer from images of an alloy of two metal mixtures. A neural network based on the U-Net architecture was implemented. Using two metrics, the effectiveness of three loss functions was investigated. Results are compared with results of manual measurement of the transition layer.

Введение

SPS-спекание – это технология обработки порошковых материалов для создания новых изделий. Порошковые смеси из металлов укладываются слоями в специальную форму и подвергаются воздействию высоких температур и электрических зарядов [1]. В результате между ними образуется переходный слой, ширина которого влияет на свойства продукта. Измерение ширины переходного слоя осуществляется вручную с помощью снимков сплавов. Автоматизация этого процесса позволит существенно ускорить исследования и повысить точность измерений.

Для автоматического измерения переходного слоя необходимо сегментировать его область на снимке сплава. За последнее десятилетие задачу автоматической сегментации преимущественно решают с помощью свёрточных нейронных сетей.

В рамках данной работы была поставлена цель – исследовать возможности применения свёрточных нейронных сетей для решения задачи сегментации переходного слоя по снимкам сплавов двух смесей металлов.

Постановка задачи

Снимки сплавов двух смесей металлов являются чёрно-белыми изображениями с размерами 1024×665 пикселей. Всего в датасете содержится 59 снимков. Каждое изображение можно разделить на три области: сталь Э76, переходный слой и сталь 12Х18Н10Т. Область переходного слоя отличается «плоской» текстурой. На рисунке 1 пример снимка сплава и соответствующей ему маски с тремя областями.

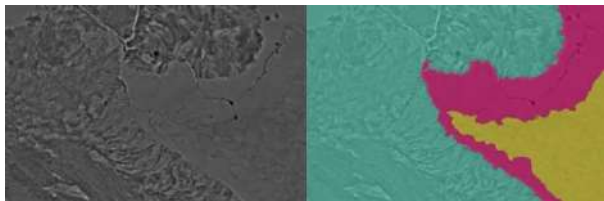


Рис. 1 – Снимок сплава двух смесей металлов и его маска.

Задача заключается в том, чтобы определить модель свёрточной нейронной сети для задачи многоклассовой сегментации, дающую наименьшую погрешность при измерении ширины переходного слоя.

Описание датасета

Датасет был разделен на три части: 40 изображений в обучающей выборке, 10 – в валидационной и 9 – в тестовой. Для каждого

изображения созданы маски с метками, указывающими принадлежность соответствующего пикселя к определённому классу.

Так как для обучения нейросети необходимо гораздо больше данных, была проведена аугментация данных. Выбраны два метода – случайный поворот не более чем на 10 градусов и случайное отражение, так как, в отличие от других методов, они не искажают интерпретацию снимков.

Разработка свёрточной нейронной сети

Для задачи сегментации изображения была выбрана наиболее распространённая для анализа изображений архитектура U-Net [2]. В качестве энкодера используется модель MobileNetV2 [3], предобученная на датасете ImageNet. В качестве оптимизатора был выбран алгоритм Adam. На вход сети подаётся изображение с размерами 224×224 пикселей. На выходе – массив размера 224×224×3 с вероятностями принадлежности соответствующего пикселя изображения к каждому классу.

Для оценки качества сегментации были выбраны две метрики:

- Categorical Accuracy (CA) - вычисляет, как часто результат нейросети совпадают с истинным:

$$CA = \frac{TP + TN}{TP + FP + TN + FN},$$

где TP – истинно-положительные результаты, TN – истинно-отрицательные, FP – ложно-положительные, FN – ложно-отрицательные;

- Intersection-Over-Union (IoU) – распространённая метрика для семантической сегментации изображения, оценивающая степень пересечения сегментированной области объекта с его истинной областью:

$$IoU = \frac{TP}{TP + FP + FN}.$$

Для исследования были выбраны следующие функции потерь:

- Categorical Crossentropy (CCE) - мера разницы между прогнозируемым и истинным распределением вероятностей:

$$CCE(y, \hat{y}) = - \sum_i y_i \cdot \ln \hat{y}_i,$$

где y – истинный результат, \hat{y} – результат работы нейросети.

- Categorical Focal Crossentropy (CFCE) – добавляет коэффициент фокуса γ , уменьшающий значимость простых примеров и фокусирующий обучение нейросети на сложных [4]:

$$CFCE(y, \hat{y}) = (1 - p)^\gamma \cdot CCE(y, \hat{y}),$$

$$\text{где } p = \begin{cases} \hat{y}, & \text{если } y = 1 \\ 1 - \hat{y}, & \text{иначе} \end{cases}.$$

- Dice – количественно оценивает сходство между предсказанной сегментацией и истинной путём вычисления коэффициента Dice:

$$Dice(y, \hat{y}) = 1 - \frac{2 \sum_i y_i \hat{y}_i + 1}{\sum_i (y_i + \hat{y}_i) + 1}.$$

Результаты

Для оценки эффективности работы нейросети была произведена серия экспериментов, в которых количество эпох обучения было равно 80.

Исследованы методы аугментации данных, при которых обучающая и валидационная выборки увеличивались в 5 раз. Применялись два варианта аугментации: первый – случайный поворот, второй – **случайный** поворот с отражением. Для обучения использовалась функция потерь CCE, а оценка качества обучения проводилась с использованием описанных выше метрик. IoU вычислялась для класса переходного слоя (Class 2) и в среднем для всех классов. Ширина переходного слоя измерялась по результатам сегментации и сравнивалась с ручными измерениями для вычисления средней погрешности δ . Результаты исследования представлены в таблице 1.

Таблица 1 – Результаты применения методов аугментации

Вариант аугментации	CA	IoU-Class 2	IoU-Mean	δ , мкм
1	91% \pm 1%	53% \pm 3%	77% \pm 1%	1.98
2	88% \pm 2%	52% \pm 2%	75% \pm 1%	3.05

Можно сделать вывод, что достаточно воспользоваться одним методом аугментации – случайным поворотом, так как при его

использовании наблюдаются лучшие показания метрик, а также достигается меньшая погрешность измерения переходного слоя.

Затем было исследовано влияние коэффициента фокуса γ функции потери CFCE. Коэффициент менялся от 0 до 4. Значения всех метрик менялись незначительно: СА от 90% до 92%, IoU-Class 2 – 51-52%, IoU-Mean – 76-77%. Однако при $\gamma = 0.5$ было получено наименьшее значение погрешности измерения переходного слоя (таблица 2).

Таблица 2 – Результаты исследования коэффициента фокуса

Значение	0.0	0.5	1.0	1.5	2.0	3.0	4.0
δ , мкм	2.14	1.7	1.9	2.0	2.3	2.07	2.17

Затем было исследовано влияние функций потерь. Значения всех метрик менялись незначительно: СА от 91% до 92%, IoU-Class 2 – 50-53%, IoU-Mean – 76-77%. Наименьшая погрешность измерения переходного слоя достигается при функции CFCE (таблица 3).

Таблица 3 – Результаты исследования функций потерь

Функция потерь	CCE	CFCE, $\gamma = 0.5$	Dice
δ , мкм	1.98	1.7	1.85

Заключение

В данной работе были исследованы возможности применения свёрточных нейронных сетей для измерения ширины переходного слоя по снимкам сплава двух смесей металлов. Нейронная сеть реализована по архитектуре U-Net и обучена на 50 снимках сплавов с использованием аугментации данных. Для оценки эффективности применения нейросети использовались такие метрики как СА и IoU. Также по результатам сегментации измерялась ширина переходного слоя и сравнивалась с результатами ручного измерения. Исследованы функции потерь, такие как CCE, CFCE и Dice. Полученные результаты свидетельствуют о возможности нейросетей сегментировать переходный слой на снимках сплавов. Необходимо дальнейшее исследование темы для повышения точности измерений по результатам сегментации с привлечением большего объема снимков для обучения нейросети.

Литература:

1. Григорьев Е.Г., Калинин Б.А. Электроимпульсная технология формирования материалов из порошков: Учебное пособие. М.: МИФИ, 2008. – 152 с.
2. U-Net: Convolutional Networks for Biomedical Image Segmentation / Olaf Ronneberger, Philipp Fischer, and Thomas Brox // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, Vol 9351. Springer, Cham. – P. 234–241. – doi: 10.1007/978-3-319-24574-4_28.
3. Sandler M., Howard A., Zhu M., Zhmoginov A. and Chen L. -C. MobileNetV2: Inverted Residuals and Linear Bottlenecks // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. – Salt Lake City, UT, USA, 2018. – P. 4510–4520. – doi: 10.1109/CVPR.2018.00474.
4. Lin T. -Y., Goyal P., Girshick R., He K. and Dollár P. Focal Loss for Dense Object Detection // IEEE Transactions on Pattern Analysis and Machine Intelligence – 2020. – Vol. 42, N 2. – P. 318–327. – doi: 10.1109/TPAMI.2018.2858826.
5. API Documentation. TensorFlow v2.16.1: сайт. – URL: https://www.tensorflow.org/api_docs (дата обращения: 01.11.2024).

МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ЭКОНОМИЧЕСКИХ ЦИКЛОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

А.А. Буксиков, А.В. Фаддеев

**Заполярный государственный университет им. Н.М. Федоровского,
г. Норильск, baait2002@gmail.com**

В данной статье рассматривается применение методов машинного обучения для моделирования и прогнозирования экономических циклов на основе временных рядов. Исследуются алгоритмы рекуррентных нейронных сетей и случайного леса, которые позволяют прогнозировать изменения экономических показателей и определять фазы экономических циклов. Проведен сравнительный анализ точности прогнозирования, предложены рекомендации для дальнейшего использования полученных результатов. Стабильность экономических циклов и их предсказуемость играют важную роль в формировании государственной и финансовой политики. В условиях высокой волатильности рынков возрастают требования к методам прогнозирования, которые могут учитывать нелинейные зависимости и большое количество факторов. В данной статье представлены результаты применения методов машинного обучения (ML) для предсказания экономических циклов.

This article discusses the application of machine learning methods for modeling and forecasting economic cycles based on time series. The algorithms of recurrent neural networks and random forests are studied, which allow forecasting changes in economic indicators and determining the phases of economic cycles. A comparative analysis of forecasting accuracy is carried out, and recommendations for further use of the obtained results are proposed. The stability of economic cycles and their predictability play an important role in the formation of public and financial policies. In conditions of high market volatility, the requirements for forecasting methods that can take into account nonlinear dependencies and a large number of factors are increasing. This article presents the results of applying machine learning (ML) methods to predict economic cycles.

Целью исследования является разработка и тестирование моделей машинного обучения для прогнозирования фаз экономического цикла на основе экономических данных, таких как валовой внутренний продукт (ВВП), уровень безработицы, индекс потребительских цен и объем розничных продаж. В работе анализируется точность прогноза и выявляются модели, наиболее эффективные в долгосрочной и краткосрочной перспективе.

Для моделирования использовались данные временных рядов с ежемесячной и ежеквартальной частотой (см. Таблица 1).

Таблица 1 — Параметры экономических показателей, используемых для прогнозирования

Показатель	Источник данных	Частота	Примечание
ВВП	Федеральная служба	Ежеквартально	Отражает общий экономический рост
Уровень безработицы	Бюро труда	Ежемесячно	Важный индикатор спада и подъема
Индекс потребительских цен	Центральный банк	Ежемесячно	Показатель инфляции
Розничные продажи	Министерство торговли	Ежемесячно	Индикатор потребительского спроса

Прогнозирование экономических показателей осуществлялось с использованием моделей ARIMA, LSTM и Random Forest. На основе временных рядов были построены модели с учетом следующих метрик точности: RMSE (среднеквадратическая ошибка) и MAPE (средняя

абсолютная ошибка). Формулы для этих показателей представлены ниже:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100$$

где y_i — истинное значение, \hat{y}_i — предсказанное значение, n — количество наблюдений.

Для моделей LSTM использовались следующие компоненты:

$$\begin{aligned} f_t &= \sigma(W_f * [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i * [h_{t-1}, x_t] + b_i) \\ C_t &= f_t * C_{t-1} + i_t * \tanh(W_c * [h_{t-1}, x_t] + b_c) \\ o_t &= \sigma(W_o * [h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh(C_t) \end{aligned}$$

Обучение проводилось на временных рядах, охватывающих период с 1980 по 2020 гг., с последующим тестированием на данных за последние 10 лет. Каждая модель оценивалась по метрикам точности RMSE и MAPE для выявления наиболее подходящего метода прогнозирования экономических циклов. Результаты приведены в Таблице 2.

Таблица 2 — Сравнение точности моделей

Модель	RMSE	MAPE (%)	Среднее отклонение	Время обучения
ARIMA	12.5	8.3	10.2	10 мин
LSTM	8.1	5.4	6.8	3 ч
Случайный лес	9.3	6.1	7.9	1 ч
VAR	10.7	7.8	9.1	15 мин

Результаты тестирования показали, что модель LSTM наиболее точно предсказывает краткосрочные изменения, тогда как модель случайного леса демонстрирует стабильные результаты на долгосрочной перспективе. Таблица 3 представляет прогнозирование значений ВВП для фазы спада.

При сравнении с базовыми моделями ARIMA и VAR модели на основе машинного обучения показали преимущество в учете нелинейных зависимостей. Это позволяет более точно выявлять фазы экономического цикла, особенно в условиях высокой волатильности рынка.

Исследование показало, что модели LSTM и случайного леса могут значительно повысить точность прогнозов экономических циклов. Эти модели продемонстрировали способность точно предсказывать как краткосрочные, так и долгосрочные изменения, что делает их полезными инструментами для экономического анализа.

**Таблица 3 — Результаты прогнозирования экономических циклов
(Пример для фазы спада)**

Время	Реальное значение ВВП	Прогноз LSTM	Прогноз ARIMA	Прогноз Random Forest
2023, 1-й кв.	5000	5100	4980	5075
2023, 2-й кв.	4950	4925	4890	4910
2023, 3-й кв.	4800	4785	4750	4795
2023, 4-й кв.	4700	4685	4600	4690

Результаты могут быть использованы для разработки рекомендаций по улучшению процесса планирования в экономических и финансовых организациях, а также для поддержки государственных решений в условиях экономической нестабильности.

В дальнейшем планируется интеграция дополнительных макроэкономических факторов и изучение ансамблевых методов, что позволит повысить устойчивость моделей к внешним экономическим шокам.

Литература:

1. *Атчаде М.Н.* Адаптивные методы прогнозирования: реализация в Excel и программе R. СПб.: СПбГЭУ, 2018. – 101 с.
2. *Лукашин Ю.П.* Адаптивные методы краткосрочного прогнозирования временных рядов. М.: Финансы и статистика, 2003. - 416 с.

3. *Shcherbakov M.V., Brebels A., Shcherbakova L.V., Tyukov A.P., Janovsky T.A., Kamaev V.A.* A Survey of Forecast Error Measures. *World Applied Sciences Journal*. – 2013. – Vol. 24. – P. 171–176. – doi: 10.5829/idosi.wasj.2013.24.itmies.80032.
4. *Hochreiter, S., & Schmidhuber, J.* Long Short-Term Memory. *Neural Computation*. – 1997. – Vol. 9(8). P. 1735–1780. – doi: 10.1162/neco.1997.9.8.1735.
5. *Breiman, L.* Random Forests. *Machine Learning*. – 2001. – Vol. 45(1). – P. 5–32. – doi: 10.1023/A:1010933404324.
6. *Hyndman R.J., Athanasopoulos G.* Forecasting: Principles and Practice. 2nd ed. OTexts, 2018. URL: <https://otexts.com/fpp2/> (дата обращения: 26.10.2024).
7. *Brockwell P.J., Davis R.A.* Introduction to Time Series and Forecasting. 3rd ed. Springer. – 2016. – 426 p. – doi: 10.1007/978-3-319-29854-2.
8. *Hastie T., Tibshirani R., Friedman J.* The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. Springer. – 2009. – 745 p. – doi: 10.1007/978-0-387-84858-7.
9. *Box G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M.* Time Series Analysis: Forecasting and Control. 5th ed. Wiley, 2015. – 712 p. – doi: 10.1002/9781118619193.
10. *Chatfield C.* The Analysis of Time Series: An Introduction. 6th ed. CRC Press. – 2003. – 352 p. – doi: 10.1201/9780203491683.

СТАТИСТИЧЕСКИЙ АНАЛИЗ ДИСТАНЦИОННОГО ОБУЧЕНИЯ

Л.О. Гафарова

Новосибирский государственный университет

экономики и управления «НИНХ»,

г. Новосибирск, e.v.makaridina@edu.nsuem.ru

Научный руководитель: Макаридина Е.В., ст. преподаватель

В статье представлен статистический анализ дистанционного обучения, основанный на итогах опроса студентов, изучающих дисциплину «Статистика» в онлайн формате. Основные факторы, влияющие на успеваемость изучении дисциплины, являются: использованное количество времени на обучение, качество материалов и связь с преподавателем. На основе корреляционного анализа была изучена взаимосвязь между успеваемостью и временем, проведённым на обучение.

The article presents a statistical analysis of distance learning based on the results of a survey of students studying the discipline "Statistics" in an online format. The main factors influencing the academic performance of the discipline are: the amount of time spent on training, the quality of materials and communication with the teacher. On the basis of correlation and regression analysis, the relationship between academic performance and time spent on training was studied.

В последние годы дистанционное обучение стало неотъемлемой частью образовательного процесса. С ускорением развития технологий, а также с учётом глобальных изменений на сегодняшний день, учебные заведения были вынуждены добавить онлайн формат изучения дисциплин в свои образовательные программы. Это открыло для студентов и преподавателей новые возможности в организации учебного процесса.

Статистический анализ дистанционного обучения позволяет понять, как такой формат влияет на успеваемость, оценивая различные факторы, например, мотивацию студентов, удовлетворённость образовательным процессом, восприятие учебного процесса, взаимодействие с преподавателем и время, которое обучающийся уделяет изучением различных дисциплин.

Так, онлайн формат помогает изучать предметы в удобное время, не зависимо от стабильного расписания и места проведения занятий. Что позволяет, в свою очередь, гибко выстраивать своё расписание изучение предметов. Появляется возможность отводить больше времени на изучение дисциплин, с которыми возникают трудности. В данной работе был проведён статистический анализ дистанционного обучения, на основе зависимости между успеваемостью и временем, которое учащиеся использовали для обучения. Корреляционный анализ позволяет выявить зависимость между ними.

Задачи корреляционного анализа сводятся к измерению тесноты связи между варьирующими признаками, определению неизвестных причинных связей и оценке факторов, оказывающих наибольшее влияние на вариацию результативного признака [1].

Перед использованием корреляционного анализа была построена таблица на основе опроса, где фактором являлось затраченное время на обучение, а результатом – успеваемость. Это позволит проанализировать данные и оценить зависимость между ними.

Для изучения зависимости между количеством часов, проведённых за обучением и успеваемости были использованы следующие показатели:

1. Коэффициент корреляции — статистический показатель, который измеряет степень зависимости между переменными. Он принимает

значение в диапазоне от -1 до 1, где -1 – полная отрицательная корреляция; 0 — отсутствие линейной корреляции; 1 – полная положительная корреляция

2. t-Критерий Стьюдента – оценка значимости коэффициента корреляции. Это позволяет определить, является ли наблюдаемая корреляция между переменными статистически значимой.

3. Коэффициент детерминации – характеризует долю дисперсии результативного признака «у», объясняемую регрессией, в общей дисперсии результативного признака. Определяет, какая доля вариации признака «у» учтена в модели и обусловлена влиянием на него факторов, включаемых в модель. Высокий R^2 говорит о том, что факторы X предсказывают большую долю изменений «у».

Источником данных является онлайн опрос обучающихся, которые находятся на дистанционном обучении [2].

Стоит отметить, что общее количество респондентов составило 80 студентов, изучающие дисциплину «Статистика» в дистанционном формате. На основе полученных данных, был произведён корреляционный анализ для выявления ключевых факторов, влияющих на эффективность данного формата обучения.

В таблице 1 представлены данные об успеваемости и количестве часов в день, проведённых за обучением.

Таблица 1 – Данные об успеваемости и количестве часов в день, проведённым на обучение

Оцен- ка за семест р	Кол-в о часов	Оцен- ка за семест р	Кол-в о часов	Оцен- ка за семест р	Кол-в о часов	Оцен- ка за семест р	Кол-в о часов
3	1	5	3	3	2	2	0
5	3	2	0	5	4	3	1
4	5	3	1	4	6	3	2
5	3	3	1	2	0	3	1
5	4	2	0	4	4	3	1
4	3	5	5	2	0	3	1
2	0	4	6	5	6	5	4
5	5	4	3	2	0	3	1
3	2	2	0	5	3	5	4
2	0	5	4	2	0	4	5
3	2	5	5	5	6	3	1
5	6	4	3	3	2	5	4
2	0	2	0	5	6	4	4

2	0	2	0	5	5	4	6
3	1	5	5	2	0	3	2
3	2	4	4	5	4	2	0
2	0	3	1	4	5	5	4
4	3	4	4	4	4	5	6
4	3	3	2	3	2	4	5
4	4	2	0	4	3	2	0

Расчёт корреляции выявил сильную статистическую связь между успеваемостью и количеством часов, проведённым на обучении, так как коэффициент корреляции составил 0,89. Критерий Стьюдента выявил, что количество часов, проведённых на обучении, значимо влияет на успеваемость. Модель показала, что 79% вариации успеваемости можно объяснить временем занятий.

Результаты показывают, что дистанционное обучение может быть эффективным при условии активного вовлечения студента и достаточного количества времени, уделяемого на учёбу. Выявленные закономерности могут помочь для оптимизации учебного процесса и улучшения методик преподавания.

Статистический анализ дистанционного обучения позволил выявить ключевые факторы, влияющие на успеваемость студентов. Результаты показали, что активное вовлечение студентов, адекватное распределение времени учебного процесса и качество взаимодействия с преподавателями являются основными факторами, способствующими успешному усвоению материала.

Анализ данных подтвердил наличие значительной положительной корреляции между временем, проведённым на обучении, и успеваемостью, что подчёркивает важность организации учебного процесса. Выявленные закономерности могут быть основой для создания рекомендаций, включая интерактивные методы и технологии обучения, которые способствуют более глубокому вовлечению студентов в учебный процесс.

Так, результаты могут подтвердить, что дистанционное обучение достаточно эффективно, при активном вовлечении студента. Важно продолжать изучение влияния различных факторов на образовательные результаты, чтобы создать улучшенные и более адаптивные образовательные среды, которые соответствуют потребностям современных студентов.

Литература:

1. Глинский В.В., Ионин В.Г., Л.К. Серга [и др.]; под ред. В.Г. Ионина. – 4-е., перераб. и доп. – Статистика: учебник. – М.: ИНФРА-М, 2017. – 355 с.

2. Опрос студентов «анализ эффективности дистанционного обучения»: [Электронный ресурс]. – URL: <https://docs.google.com/forms/d/1jgbFB4tzmFC2aeW9upn0vIx2sBtUAAq6mlAP-5qCTP4/viewform> (дата обращения: 03.11.2024).

ОСОБЕННОСТИ ИСПОЛЬЗОВАНИЯ КЛАСИЧЕСКИХ МЕТОДОВ И НЕЙРОННЫХ СЕТЕЙ В ПРОГНОЗИРОВАНИИ ВРЕМЕННЫХ РЯДОВ

В.А. Громов, А.В. Фаддеевков

**Заполярный государственный университет им. Н.М. Федоровского,
г. Норильск, хр34gvova@gmail.com**

Статья посвящена методам прогнозирования временных рядов, сгруппированным на три категории: классические статистические методы, методы машинного обучения и глубокое обучение. Классические методы подходят для небольших данных с заданными свойствами, такими как стационарность и сезонность. Методы машинного обучения требуют большего объема данных и ресурсов, но обеспечивают гибкость в работе с нелинейными зависимостями. Глубокое обучение позволяет выявлять сложные закономерности и достигать высокой точности прогнозов на больших данных.

The paper is devoted to time series forecasting methods grouped into three categories: classical statistical methods, machine learning methods and deep learning. Classical methods are suitable for small data with specified properties such as stationarity and seasonality. Machine learning methods require more data and resources, but provide the flexibility to deal with non-linear dependencies. Deep learning can identify complex patterns and achieve high prediction accuracy on big data.

На данный момент существует множество различных методов прогнозирования временных рядов, каждый из которых имеет свои отличительные особенности и показывает разную степень эффективности в зависимости от вводных. На данный момент их можно сгруппировать как:

1. Классические статистические методы
2. Методы машинного обучения
3. Глубокое обучение (Deep Learning)

Классические методы прогнозирования лучше всего показывают свою эффективность на небольшом объеме данных с заранее указанными свойствами. К таким свойствам относят стационарность (статистические характеристики данных не изменяются со временем), линейность зависимостей, сезонность, наличие “шумов” и автокорреляция (взаимосвязь значений временного ряда с предыдущими наблюдениями). Представителями классические статистические методов являются:

1. ARIMA (AutoRegressive Integrated Moving Average) – это один из самых популярных методов для прогнозирования стационарных временных рядов [1]. Метод основывается на трех компонентах: авторегрессии (AR), скользящей средней (MA) и интегрировании (I), что позволяет моделировать широкий спектр временных зависимостей. Однако он не подходит для данных с явной нелинейной структурой.
2. SARIMA (Seasonal ARIMA) – модификация ARIMA с учетом сезонности [2]
3. VAR (Vector AutoRegression) - многомерный метод, который учитывает взаимосвязи нескольких временных рядов [3]. Применяется, когда необходимо моделировать систему взаимозависимых временных рядов, таких как экономические показатели.
4. GARCH (Generalized Autoregressive Conditional Heteroskedasticity) – используется для моделирования временных рядов с переменной волатильностью. Особенно эффективен для анализа и прогнозирования финансовых данных [4]

Методы машинного обучения кардинально отличаются от классических методов прогнозирования. Они менее требовательны к предварительной подготовке и анализу данных, но нуждаются в большем объеме информации для обучения. Также методы машинного обучения потребляют значительно более вычислительных ресурсов, но взамен предоставляют гибкость в работе с нелинейными зависимостями и сложными данными. Представителями методов машинного обучения, применяемых для задач прогнозирования временных рядов, можно отнести:

1. Linear, Ridge, Lasso – линейные модели, хорошо работающие на высокоразмерных данных [5].
2. Random Forest, Gradient Boosting, XGBoost, LightGBM – ансамблевые методы, которые эффективно извлекают нелинейные зависимости [6].

3. SVM (Support Vector Machines), k-NN (k-Nearest Neighbors) – методы, подходящие для временных рядов с разной степенью сложности [7].

Методы глубокого обучения по своей сути являются частью машинного обучения, но в данном случае выделяется как отдельная категория из-за особенностей архитектуры и сложности реализации.

Глубокое обучение основывается на многослойных искусственных нейронных сетях. Оно требует больших временных и вычислительных затрат на подбор архитектуры с последующим обучением, но взамен предоставляет возможность автоматического выявления сложных нелинейных закономерностей. Это позволяет составлять наиболее точные прогнозы, особенно на основе больших массивов данных. К методам глубокого обучения можно отнести:

1. RNN (рекуррентные нейронные сети) - архитектура, разработанная для работы с последовательностями данных.
2. LSTM (долгосрочную краткосрочную память) - усовершенствованная версия RNN, предназначенная для работы с длинными последовательностями.
3. GRU (ускоренные версии LSTM) - модификация LSTM с упрощенной архитектурой [8].
4. Transformer – мощная архитектура, изначально разработанная для обработки текстовых данных [9]. Transformers успешно используются для временных рядов благодаря механизму «внимания», который выделяет ключевые зависимости в данных независимо от их длины.

Глубокое обучение становится незаменимым для задач, связанных с большими объемами данных, такими как анализ временных рядов. Несмотря на высокую сложность реализации, оно предоставляет впечатляющие результаты там, где другие методы ограничены.

Литература:

1. Box G.E.P., Jenkins G.M., Reinsel G.C., & Ljung G.M. Time Series Analysis: Forecasting and Control. John Wiley & Sons, 2015. – 712 с.
2. Hyndman R.J., & Athanasopoulos G. Forecasting: Principles and Practice (2nd ed.), 2018. – 382 с.
3. Lütkepohl H. New Introduction to Multiple Time Series Analysis. Springer, 2005. – 765 с.
4. Engle R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation. *Econometrica*, – 1982. – Vol. 50(4). – P. 987–1007. – 20 с.

5. *Hastie T., Tibshirani R., & Friedman J.* The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009. – 764 с.
6. *Friedman J.H.* Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics. – 2001. – Vol. 29(5). – P. 1189–1232. – 46 с.
7. *Cortes C., & Vapnik V.* Support-vector networks. Machine Learning. – 1995. – Vol. 20. – P. 273–297. – 25 с.
8. *Cho K., Van Merriënboer B., Gulcehre C., Bahdanau D., Bougares F., Schwenk H., & Bengio Y.* Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv preprint arXiv:1406.1078. – 2014. – 11 с.
9. *Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., & Polosukhin I.* Attention Is All You Need. Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 5998–6008. – 11 с.

ОЦЕНКА ПАРАМЕТРОВ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ МЕТОДОМ Lq-ПРАВДОПОДОБИЯ

В.К. Дворецкая

**Новосибирский государственный технический университет,
г. Новосибирск, dvoreczkaya.2018@stud.nstu.ru**

Научный руководитель: Тимофеев В.С., д.т.н., доцент

Получена оценка параметров нормального распределения с помощью метода максимального правдоподобия и метода Lq-правдоподобия. Проведено исследование на смоделированных данных, описывающих нормальное распределение.

An estimate of the parameters of the normal distribution was obtained using the maximum likelihood method and the Lq likelihood method. A study was conducted on simulated data describing the normal distribution.

В статистическом анализе оценка параметров распределений играет ключевую роль в понимании и интерпретации данных. Поэтому правильный выбор метода оценки становится критически важным, особенно учитывая разнообразие типов распределений. Однако на практике далеко не все распределения встречаются одинаково часто. Нормальный закон распределения или закон Гаусса играет важную роль в статистике и занимает особое положение среди других законов [1]. Нормальный закон распределения характеризуется плотностью вероятности вида:

$$f(x; \mu; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (1)$$

где $x \in (-\infty; +\infty)$ $x \in \{x_1, x_2, \dots, x_n\}$, μ - математическое ожидание, σ - среднее квадратическое отклонение.

Стандартным подходом для оценки параметров распределения является метод максимального правдоподобия (ММП). Этот метод основан на максимизации функции правдоподобия, которая описывает вероятность наблюдения данных при заданных параметрах [2]. Для выборки из n независимых наблюдений x_1, x_2, \dots, x_n функция правдоподобия нормального распределения (1) имеет вид:

$$L(x_1, x_2, \dots, x_n; \mu; \sigma) = \prod_{i=1}^n \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \right). \quad (2)$$

Часто удобно иметь дело не с функцией (2), а ее логарифмом. Логарифмирование функции правдоподобия и поиск оценок параметров нормального распределения приведен в [2]. В данной работе будут использоваться сами оценки $\hat{\mu}$ и $\hat{\sigma}$, которые имеют вид:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (3)$$

и

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2}. \quad (4)$$

ММП является мощным инструментом для оценки параметров распределений. Однако при анализе небольшой выборки с помощью ММП, полученные результаты могут иметь большую дисперсию, что увеличивает риск ошибочных выводов в исследовании и снижает достоверность результатов. В связи с этим более корректным представляется использование альтернативных методов оценки. Интересным представителем таких методов является метод Lq-правдоподобия (MLqП), предложенный в [3].

Отличительной особенностью MLqП является то, что вместо функции правдоподобия используется L_q -функция, которая определяется как:

$$L_q(x_1, x_2, \dots, x_n; \mu; \sigma) = \begin{cases} \sum_{i=1}^n \ln f(x_i; \mu; \sigma), & \text{если } q = 1 \\ \frac{1}{1-q} \left(\sum_{i=1}^n f(x_i; \mu; \sigma)^{1-q} - 1 \right), & \text{иначе} \end{cases}, (5)$$

где q - параметр, контролирующий форму функции (параметр искажения). При $q = 1$ L_q -функция соответствует логарифму функции правдоподобия исследуемого распределения и MLqП оценки совпадают с оценками ММП.

Параметр q в MLqП может принимать значения от 0 до 2 [4]. Если $q < 1$ MLqП акцентирует внимание на маловероятных событиях, что позволяет более точно оценивать параметры распределений в присутствии выбросов. При $q > 1$ снижается чувствительность к выбросам. Определение корректного q влияет на точность оценки распределения.

Проведем вычисления позволяющие получить оценку параметров нормального распределения с помощью MLqП, вычисляя логарифм L_q -функции (5) при $q \neq 1$:

$$\ln L_q(x_1, x_2, \dots, x_n; \mu; \sigma) = \ln \left(- \sum_{i=1}^n \frac{f(x_i; \mu; \sigma)^{1-q} - 1}{1-q} \right),$$

$$\ln L_q(x_1, x_2, \dots, x_n; \mu; \sigma; q) = \ln \left(- \sum_{i=1}^n \frac{f(x_1, x_2, \dots, x_n; \mu; \sigma)^{1-q} - 1}{1-q} \right)$$

тогда

$$\ln L_q(x_1, x_2, \dots, x_n; \mu; \sigma) = - \sum_{i=1}^n (1-q) \left(-n \ln \sigma - \frac{n}{2} \ln \sqrt{2\pi} - \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right) + n \ln(1-q).$$

$$\ln L_q(u) = - \sum_{i=0}^n (1-q) \left(-n \ln \sigma - \frac{n}{2} \ln \sqrt{2\pi} - \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right) + n \ln(1-q).$$

В результате взятия производной $\ln L_q(x_1, x_2, \dots, x_n; \mu; \sigma)$ по каждому из параметров $\ln L_q(x_1, x_2, \dots, x_n; \mu; \sigma)$, оценка $\hat{\mu}$ будет равна (3), а оценки $\hat{\sigma}$ и \hat{q} будут определены как:

$$\hat{q} = 1 - e^{-1}, (6)$$

$$\hat{\sigma} = \sqrt{\frac{1}{n\hat{q}} \sum_{i=1}^n (x_i - \hat{\mu})^2}. (7)$$

Найденные оценки (3), (4), (6) и (7), функция правдоподобия (2) и L_q -функция (5) для нормального распределения реализованы на языке программирования Python. Для исследования были смоделированы выборки нормального распределения с n от 100 до 10000 при разных параметрах μ и σ . Для сравнения методов оценки была введена относительная эффективность R_μ и R_σ :

$$R_\mu = MSE(\hat{\mu})/MSE(\bar{\mu}) \quad (8)$$

$$R_\sigma = MSE(\hat{\sigma})/MSE(\bar{\sigma}) \quad (9)$$

где

$$MSE(\hat{\mu}) = \frac{1}{m} \sum_{i=1}^m (\mu - \hat{\mu}_i)^2, \quad (10)$$

$$MSE(\hat{\sigma}) = \frac{1}{m} \sum_{i=1}^m (\sigma - \hat{\sigma}_i)^2, \quad (11)$$

$\mu, \sigma, y_i \in \{\mu, \sigma\}$ - фактическое значение параметра, $\hat{\mu}_i, \hat{\sigma}_i, \hat{\mu}, \hat{\sigma} \in \{\hat{\mu}, \hat{\sigma}\}$ - оцененные i -е параметры, $m = 500$ - количество смоделированных выборок для определенной размерности, в соответствии с (10) и (11),

$MSE(\hat{\mu})$ и $MSE(\hat{\sigma})$ квадратичные ошибки оцененных параметров ММП, $MSE(\bar{\mu})$ и $MSE(\bar{\sigma})$ квадратичные ошибки оцененных параметров MLqП.

Если $R_\sigma < 1$ и $R_\mu < 1$, $R_\sigma \rightarrow 0$ и $R_\mu \rightarrow 0$ это указывает на то, что метод ММП является более предпочтительным по сравнению с методом MLqП. Рассмотрим пример стандартного нормального распределения на рисунке 1.

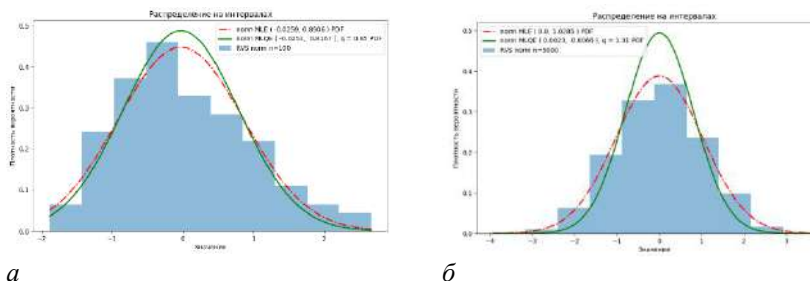


Рис. 1 – Нормальное распределение с ожидаемые параметры $\mu = 0, \sigma = 1$, а: $n = 100$, б: $n = 5000$, $q \in [0.85; 1.53]$

Для а: ММП $\hat{\mu} = -0.0259$, $\hat{\sigma} = 0.8906$ (сплошная линия), MLqП $\hat{\mu} = -0.0252$, $\hat{\sigma} = 0.8167$, $\hat{q} = 0.85$ (штрихпунктирная линия) и относительная эффективность $R_{\mu} = 1.0535$ и $R_{\sigma} = 0.6561$. В случае оценки $\hat{\sigma}$, R_{σ} меньше 1, однако, не выполняется условие $R_{\sigma} \rightarrow 0$ и $R_{\mu} \rightarrow 0$, поэтому MLqП показал хорошие результаты в оценке параметров для $n = 100$. Для б: ММП $\hat{\mu} = -0.0235$, $\hat{\sigma} = 1.0295$ (сплошная линия), MLqП $\hat{\mu} = -0.0232$, $\hat{\sigma} = 0.7970$, $\hat{q} = 1.31$ (штрихпунктирная линия) и относительная эффективность $R_{\mu} = 0.9282$, $R_{\sigma} = 0.0021$ меньше 1, в данном случае $R_{\sigma} \rightarrow 0$, поэтому ММП имеет преимущество в оценке параметра σ для

$$L(x_1, x_2, \dots, x_n | \mu, \sigma) = \prod_{i=1}^n \left(\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \right).$$

$n = 5000$.

На текущий момент в научной литературе существует множество исследований, посвященных оценке параметров распределений Парето, экспоненциального и стандартного гамма-распределения с помощью MLqП [4-6]. В дальнейшем планируется провести исследования и проверить утверждения из [4-6] на смоделированных данных, описывающих нормальное распределение.

Литература:

1. Вентцель Е.С. Теория вероятностей: Учеб. для вузов. – 10-е изд. – М.: Высш. шк., 2006. – 575 с.
2. Клепиков Н.П., Соколов С.Н. Анализ и планирование экспериментов методом максимума правдоподобия. – М.: «Наука», 1964. – 184 с.

3. Ferrari Davide, Yang Yuhong. Maximum Lq-Likelihood Estimation // The Annals of Statistics. – 2010. – Vol. 38, No. 2. – P. 753–783.
4. Jingjing Wu, Nana Xing, Shawn Liu. Maximum Lq-likelihood Estimation for Gamma Distributions // Journal of Advanced Statistics. – 2017. – Vol. 2, No. 1. – P. 54-70.
5. Nana Xing. Maximum Lq-Likelihood Estimation for Gamma Distributions: Graduate program in Mathematics and Statistics: Calgary, Alberta, 2015. – 65 p.
6. Ferrari D., Paterlini S. The maximum lq-likelihood method: an application to extreme quantile estimation in finance, Methodology and Computing in Applied Probability. – 2007. – Vol. 11, No. 1. – P. 3–19.

СЕГМЕНТАЦИЯ X-RAY ИЗОБРАЖЕНИЙ ПОЗВОНОЧНИКА С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

А.А. Ерошенко

**Новосибирский государственный технический университет,
г. Новосибирск, artiom.inety@gmail.com**

Научный руководитель: Тимофеев В.С., д.т.н., доцент

В данной работе предлагается решение задачи расчета координат реперных точек позвонков (от S1 до C2) по X-Ray изображению для помощи врачам-вертебрологам в дальнейшей постановке диагноза. Разрабатывая полностью автоматическую модель, предлагается использовать нейронную сеть U-Net++ архитектуры для сегментации позвоночника на X-Ray снимке, а также алгоритм нахождения реперных точек позвонков, основанный на множественных аппроксимациях контура позвоночника и скользящей нормали.

This paper proposes a solution to the problem of calculating coordinates of reference points of vertebrae (from S1 to C2) from an X-Ray image to help vertebrologists in further diagnosis. Developing a fully automatic model, we suggest using the U-Net++ neural network for segmenting the spine on an X-Ray image, as well as an algorithm for finding coordinates of reference points of vertebrae based on multiple approximations of the spine contour and a sliding normal.

Существует множество методов современной диагностики состояния позвоночника человека, но в клинических случаях в основном преобладает анализ X-Ray изображений, где происходит обработка размеров, объемов и других необходимых параметров позвонков. На данный момент такая обработка производится вручную, что сильно снижает скорость постановки диагноза, а также требует предельной внимательности и профессионализма от специалистов. В связи с чем и была поставлена задача создания программного решения для

автоматического выделения позвонков (от S1 до C2) на X-Ray изображениях и расчета координат их реперных точек.

Для решения задачи выделения позвонков на X-Ray изображении в данной работе предлагается использовать нейронную сеть U-Net++ [1] архитектуры.

Данная сеть состоит из трёх частей: сужающейся (энкодер), расширяющейся (декодер) и внутренней. Энкодер состоит из основных блоков, включающих две операции свертки, за каждой из которых следует операция активации. За каждым блоком следует операция подвыборки по максимальному значению. Декодер и внутренние блоки состоят из шагов обратной свертки, которая уменьшает количество каналов, после чего следует конкатенация с картой признаков от соответствующего предыдущего слоя и основной блок. Такой подход позволяет сократить семантический разрыв между картами признаков кодера и декодера и, как следствие, повысить точность сегментации. Кроме того, в данной архитектуре реализовано «глубокое наблюдение» (Deep supervision), которое позволяет настраивать сложность модели, отсекая слои, и обучать эти слои по отдельности.

В качестве функции активации используется ELU

$$ELU(x) = \begin{cases} x, & x > 0 \\ \alpha(e^x - 1), & x \leq 0 < \alpha \end{cases}.$$

В качестве метрики для оценки качества сегментации был выбран коэффициент Дайса [4]

$$Dice(X, \hat{X}) = \frac{2X\hat{X}}{X + \hat{X}}, \quad (1)$$

где X - исходные данные, а \hat{X} - результат работы нейронной сети.

Функция потерь [5] определяется через объединение бинарной перекрестной энтропии [8] и коэффициента Дайса (1):

$$\begin{aligned} loss(X, \hat{X}) &= 2 - (Crossentropy(X, \hat{X}) + Dice(X, \hat{X})), \\ Crossentropy(X, \hat{X}) &= \hat{X} \log_2(X) + (1 - \hat{X}) \log_2(1 - X). \end{aligned}$$

Имея нейронную сеть, которая выделяет позвоночник на X-Ray изображении, можно переходить к расчету координат реперных точек позвонков.

Для решения данной задачи предлагается следующий алгоритм:

1. Используя квантильную регрессию [6], аппроксимируем левый и правый контур позвоночника полиномом Бернштейна [7]

$$P(x) = c_0 + c_1\varphi_1(x) + c_2\varphi_2(x) + \dots + c_M\varphi_M(x), \quad x \in \delta_{spine}. \quad (2)$$

$$C = (c_0, c_1, \dots, c_M)^T = (A^T A)^{-1} (A^T Y),$$

где $Y = (y_1, \dots, y_N)^T$ - контур позвоночника,

$$A = W \begin{pmatrix} 1 & \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_M(x_1) \\ 1 & \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_M(x_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \varphi_1(x_N) & \varphi_2(x_N) & \cdots & \varphi_M(x_N) \end{pmatrix},$$

$$\varphi_i(x) = \left(\frac{x - x_1}{x_N - x_1} \right)^i \left(1 - \frac{x - x_1}{x_N - x_1} \right)^{M-i},$$

$$W = \text{diag}(\rho_\tau((y_1 - A_1 C)^2), \dots, \rho_\tau((y_N - A_N C)^2)), \tau \in [0, 1].$$

$$\tau_{\text{left}} = \arg \min_{\tau_{\text{left}}} (1 - \gamma(X_{\tau_{\text{left}}, \tau_{\text{right}}}, X)),$$

$$\tau_{\text{right}} = \arg \min_{\tau_{\text{right}}} (1 - \gamma(X_{\tau_{\text{left}}, \tau_{\text{right}}}, X)),$$

где $X_{\tau_{\text{left}}, \tau_{\text{right}}}$ - позвоночник, ограниченный P_{left} и P_{right} ,

$$\gamma(X, Y) = \frac{X \cap Y}{X \cup Y}.$$

2. Задаем полином $P_{\text{mid}}(x)$ с коэффициентами

$$c_{i-\text{mid}} = \frac{c_{i-\text{left}} + c_{i-\text{right}}}{2}, i = \overline{0, M}.$$

3. Строим нормаль

$$f(x, u_i) = - \left(\frac{dP_{\text{mid}}}{dx}(u_i) \right)^{-1} (x - u_i) + P_{\text{mid}}(u_i)$$

к касательной в каждой точке $u_i \in \delta_{\text{spine}}$, где

$$|x - u_i| \leq 3 \max\{|u_i - x_{\text{left}}|, |u_i - x_{\text{right}}|\},$$

при этом

$$f(x_{\text{left}}, u_i) = P_{\text{left}}(x_{\text{left}}), f(x_{\text{right}}, u_i) = P_{\text{right}}(x_{\text{right}}).$$

4. Вычисляем фактическое пороговое значение

$$\Delta = \frac{|Q|}{|Q| + |\overline{Q}|},$$

где $Q = \{(x, f(x, u_i)) | (x, f(x, u_i)) - \text{точка позвонка}\}.$

5. Положим Δ_0 и $\Delta_{N+1} = 0$. Если $\Delta_i \geq \varepsilon$ и $\Delta_j < \varepsilon$, где $j = i \pm 1, i = \overline{1, N}$, то реперные точки лежат на $f(x, u_i)$.

$$\varepsilon = \arg \min_{\varepsilon} (1 - \gamma(X_{\varepsilon}, X)),$$

где X_{ε} - позвоночник, ограниченный последовательными парами $(f(x, u_i), f(x, u_{i+w}))$, которые удовлетворяют данному шагу.

6. Для каждой последовательной пары $(f(x, u_i), f(x, u_{i+w}))$, удовлетворяющей шагу 5, используя квантильную регрессию, аппроксимируем левый и правый контур позвонка, ограниченного данной парой, полиномом Бернштейна (2) ($M = 1$). Тогда

$$\{x_1, P_{i-left}(x_1)\}, \{x_2, P_{i-right}(x_2)\}, \{x_3, P_{i-left}(x_3)\}, \{x_4, P_{i-right}(x_4)\}$$

при том, что

$$f(x_1, u_i) = P_{i-left}(x_1), f(x_2, u_i) = P_{i-right}(x_2), f(x_3, u_{i+w}) = P_{i-left}(x_3), \\ f(x_4, u_{i+w}) = P_{i-right}(x_4),$$

являются реперными точками.

В качестве метрики для оценки точности определения координат реперных точек было выбрано простое расстояние

$$d_i(X, \hat{X}) = \sqrt{\left(\arg_x R_i(X) - \arg_x R_i(\hat{X})\right)^2 + \left(\arg_y R_i(X) - \arg_y R_i(\hat{X})\right)^2},$$

где $R_i(X)$ – i -ая реперная точка.

Для обучения и валидации вышеописанной модели были использованы анонимные изображения сагиттальной проекции позвоночника 746 пациентов в формате dicom с разрешением 1920 на 4608 пикселей. Проведением аугментации изображений поворотом на $-7^\circ, -5^\circ, -3^\circ, 3^\circ, 5^\circ, 7^\circ$ и отбором тех, которые содержали позвоночник полностью, обучающая выборка была увеличена до 4564 изображений, из них на 4008 будет производиться обучение, 556 – валидационные.

В таблице 1 представлены результаты работы модели с числом членов в полиноме Бернштейна $M = 10$, полученные на обучающей и валидационной выборках. Из них видно, что средняя ошибка в расчёте координат реперных точек позвонков оказалась меньше 2 мм.

Таблица 1 – Результаты работы модели ($M = 10$)

Метрика	Dice	d_{avg} , мм
Обучающая выборка	0.9375	1.4867
Валидационная выборка	0.8162	1.7814

Таким образом в данной работе было проведено обучение модели U-Net++ для решения задачи сегментации позвоночника человека на X-Ray изображениях. Предложен метод выделения реперных точек позвонков на основе множественных аппроксимаций контура позвоночника и скользящей нормали. Проведены тестирования модели и выполнена оценка ее точности.

Литература

1. Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, Jianming Liang. UNet++: A Nested U-Net Architecture for Medical Image Segmentation [Electronic resource] // Arxiv. – 2018, July. – URL: <https://arxiv.org/pdf/1807.10165> (accessed 06.09.2024).
2. Гладков А.В., Комиссаров В.В. Математическое моделирование опорно-двигательного аппарата / Новосиб. гос. техн. ун-т. – Новосибирск, 2023. – 114 с.
3. Гладков А.В., Комиссаров В.В. Особенности математического трехмерного моделирования позвоночника / Новосиб. гос. техн. ун-т. – Новосибирск, 2024. – 76 с.
4. Dice Lee R. Measures of the Amount of Ecologic Association Between Species // Ecology. – 1945. – Vol. 26, N 3. – P. 297-302.
5. Klebanov L., Rachev S.T., Fabozzi F. Robust and Non-Robust Models in Statistics: Nova Scientific Publishers, New York, 2009. – 317 p.
6. Koenker R., Hallock K.F. Quantile Regression // Journal of Economic Perspectives. – 2001. – Vol. 15, N 4. – P. 143-156.
7. Бернштейн С.Н. Собрание сочинений. Т. 1. Конструктивная теория функций – М.: АН СССР, 1952. – 581 с.
8. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction: Stanford University, Stanford, CA, 2009. – 764 p.

ОЦЕНКА КАЧЕСТВА КОНТЕНТА В ДИСТАНЦИОННЫХ ПРОГРАММАХ

А.Д. Железнякова

Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, e.v.makaridina@edu.nsuem.ru

Научный руководитель: Макаридина Е.В., ст. преподаватель

В статье будет рассмотрена взаимосвязь между интерактивной наполненностью курса и успеваемостью студентов с помощью корреляционного анализа. Наличие или отсутствие связи будет оцениваться с помощью следующих показателей: коэффициента корреляции, t-критерия Стьюдента, коэффициента детерминации.

The article will examine the relationship between the interactive content of the course and student performance using correlation analysis. The presence or absence of a relationship will be assessed using the following indicators: correlation coefficient, Student's t-test, coefficient of determination.

Пандемия COVID-19 показала необходимость внедрения в классическую систему образования дистанционного обучения. Активное развитие дистанционного формата обучения пришлось именно на этот период. Онлайн-формат обучения позволил за короткий промежуток времени получить полезные профессиональные навыки. Также дистанционный формат обучения позволяет учащимся самостоятельно изучить весь необходимый материал и усовершенствовать свои знания. Однако влияние качества онлайн-курсов на успеваемость обучающихся представляет серьезный вопрос.

Качество контента в дистанционном обучении напрямую влияет на эффективность образовательного процесса. Низкое качество материалов может привести к снижению мотивации студентов, ухудшению усвоения знаний и, как следствие, к снижению общего уровня образования. В условиях, когда студенты часто обучаются самостоятельно, без непосредственного взаимодействия с преподавателями, качество контента становится особенно критичным. Оно должно быть не только актуальным и научно обоснованным, но и доступным, понятным и интерактивным, чтобы поддерживать интерес и вовлеченность обучающихся.

Кроме того, с ростом числа онлайн-курсов и программ возникает необходимость в стандартизации критериев оценки качества контента. На сегодняшний день существует множество подходов и методик, однако их разнообразие затрудняет выбор наиболее эффективных

инструментов для оценки. Это создает риск появления некачественного контента на образовательном рынке, что может негативно сказаться на репутации учебных заведений и доверии со стороны студентов.

Таким образом, актуальность проблемы оценки качества контента в дистанционном обучении обусловлена не только необходимостью обеспечения высокого уровня образования, но и стремлением к созданию эффективных и конкурентоспособных образовательных программ. В условиях быстро меняющегося мира, где технологии и методы обучения постоянно эволюционируют, важно разработать и внедрить надежные механизмы оценки, которые помогут обеспечить соответствие контента современным требованиям и ожиданиям обучающихся.

Интерактивные материалы играют ключевую роль в дистанционном обучении, значительно влияя на восприятие и оценку качества контента. Наполненность курса интерактивными элементами, такими как тесты, задания и мультимедийные ресурсы, может оказывать несколько положительных эффектов на образовательный процесс и его результаты [1].

Также, интерактивные элементы помогают студентам лучше усваивать информацию. Например, использование тестов и квизов позволяет обучающимся проверять свои знания и получать мгновенную обратную связь. Это способствует более глубокому пониманию материала и запоминанию ключевых концепций. Качественный контент, который включает интерактивные задания, обычно оценивается выше, так как он способствует активному обучению.

Именно поэтому, чтобы оценить качество контента в дистанционном обучении, будем ориентироваться на наличие или отсутствие интерактивной наполняемости курса.

Корреляционный анализ – это статистический метод изучения взаимосвязи между двумя и более случайными величинами [2]. Для корреляционного анализа необходимо, в первую очередь, провести расчет показателя коэффициента корреляции. Этот показатель принимает значения от -1 до +1. Знак коэффициента определяет направление связи, а значение – силу. При отрицательном значении связь обратная, при положительном – прямая. Теснота связи представлена в таблице 1.

Для расчета показателей была взята группа студентов СТ101, которые проходили курсы по статистике на различных платформах, наличие или отсутствие интерактивного наполнения курса и их средний балл. За наличие интерактивного наполнения курса берется 1, за отсутствие – 0. Все показатели представлены в таблице 2.

Таблица 1 – Классификация коэффициентов корреляции по силе

Диапазон значений $ r $	Линейная корреляционная зависимость Y от X
До 0,2	Очень слабая
0,2-0,5	умеренная
0,5-0,7	заметная
0,7-0,9	сильная
0,9-0,99	очень сильная

Для определения наличия связи определим, что H_0 – гипотеза о том, что связь между наличием интерактивного наполнения курса и успеваемостью студента есть. Гипотеза H_1 будет говорить об отсутствии связи между двумя показателями.

Таблица 2 – Данные о наличии интерактивной наполненностью курса и средней оценкой за курс

№ платформы для обучения	Наличие интерактивного наполнения курса	Средняя оценка за курс
1	1	4,50
2	1	3,97
3	0	3,50
4	0	3,00
5	0	3,62
6	1	4,11
7	1	3,85
8	1	4,00
9	1	4,08
10	1	4,81
11	0	2,90
12	1	4,14
13	0	3,37
14	0	2,90
15	1	4,30

Проведем расчет показателей значения коэффициента корреляции. Воспользуемся формулой 1:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \times (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \times \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

где r_{xy} – коэффициент корреляции;

n – объем выборки;

x_i – i -тый элемент выборки X ;

y_i – i -тый элемент выборки Y ;

\bar{x} – среднее значение выборки X ;

\bar{y} – среднее значение выборки.

Так, коэффициент корреляции составил 0,86. Связь между наличием интерактивного наполнения курса и успеваемостью студентов прямая, сильная. t -критерий Стьюдента расчетный больше табличного значения, при расчете коэффициента детерминации мы получили значение 0,73. Это показывает, что наличие интерактивных материалов в курсе влияет на успеваемость студентов на 73%. Гипотеза H_0 – подтвердилась.

Таким образом, оценка качества контента в дистанционных образовательных программах является ключевым аспектом, определяющим эффективность и успешность обучения в условиях современного образовательного процесса. В ходе исследования было выявлено, что качество контента напрямую влияет на мотивацию студентов, их вовлеченность и уровень усвоения знаний. Внедрение систематического подхода к оценке качества контента позволит повысить уровень образования, сделать его более доступным и эффективным, а также подготовить студентов к вызовам современного мира.

Литература:

1. Токарев Г.С., Новикова С.Е. Оценка долговременных последствий введения дистанционного обучения в системе общего образования. – 2023. – С.65-80.
2. Шаныгин С.И. Корреляционный и регрессионный анализ: учебник для вузов. – 2024. – С.42.

СТАТИСТИЧЕСКИЕ ПОДХОДЫ К ОЦЕНКЕ ИЗБЫТОЧНОЙ СМЕРТНОСТИ ИЗ-ЗА COVID-19 НА ПРИМЕРЕ НОВОСИБИРСКОЙ ОБЛАСТИ

А.А. Ицкович

**Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, itskalisa@gmail.com**

Научный руководитель: Овечкина Н.И., к. э. н., доцент

COVID-19 стал одним из основных объектов исследования за последние годы. Популярным направлением исследования является изучение последствий пандемии, в том числе изучение смертности. Избыточная смертность – показатель, способный наиболее объективно оценить, в какой мере смертельной оказалась инфекция, и в настоящее время разрабатываются различные методики для расчета данного показателя. В работе применяются некоторые из методик на примере данных по Новосибирской области за 2020 и 2021 года.

COVID-19 has become a major focus of research in recent years. A popular area of research is the study of the consequences of the pandemic, including the study of mortality. Excess mortality is the indicator that can most objectively assess the extent to which the infection was fatal, and various methodologies are being developed to calculate this indicator. This paper applies some of the methodologies using the example of data for the Novosibirsk Region for 2020 and 2021.

COVID-19 в 2020 году потряс общество: кроме страха заразиться новым вирусом, люди столкнулись с непривычным для себя укладом жизни. Каждой страной и отдельными ее регионами были введены перечни мер по регулированию пандемии. Чтобы оценить масштабы последствий коронавирусной инфекции и эффективность предпринятых мер, рассчитывают избыточную смертность.

Избыточная смертность, связанная с COVID-19, определяется Всемирной организацией здравоохранения как разница между зарегистрированным числом смертей в 2020-2021 годах и числом смертей, которое ожидалось при отсутствии пандемии на основании данных предшествующих лет [1].

Число смертельных исходов от коронавируса не дает полного представления о масштабе последствий пандемии, поскольку в эти данные не входят незарегистрированные случаи. Также эти данные не учитывают сопутствующие смерти, которые были вызваны, например, перегрузкой системы здравоохранения. Чтобы получить более полную картину о влиянии пандемии на число смертей, рассчитывают избыточную смертность [2].

Расчет избыточной смертности – далеко не тривиальная задача. Существует несколько способов расчета этого показателя, и продолжают развиваться универсальные статистические методы. Первый рассматриваемый метод расчета напрямую дублирует определение избыточной смертности: необходимо вычесть из зарегистрированного числа смертей ожидаемое число смертей. В данном случае за ожидаемое число смертей могут быть приняты средние данные по смертности за 4-5 лет до ковида на данной территории [2]. При расчете избыточной смертности в Новосибирской области было взято среднее значение за 2016-2019 года, оно составило 12,5 чел. на 1000 населения. В 2020 году фактическое значение составило 14,6 чел. на 1000 населения, в 2021 году - 16,7 чел. на 1000 населения [5]. Таким образом, избыточная смертность в 2020 году в Новосибирской области составила 2,1 человека, в 2021 году - 4,2 чел. на 1000 населения.

Существует Р-показатель (или Р-оценка) для расчета избыточной смертности, он более сопоставим по регионам страны или по странам [2]:

$$P - \text{оценка} = \frac{\text{Зарегистр. смерти} - \text{Прогнозир. смерти}}{\text{Прогнозир. смерти}} \cdot 100\%$$

Для расчета данного показателя для Новосибирской области необходимо определить прогнозируемые смерти на 2020 и 2021 года. Уравнение тренда по данным числа умерших на 1000 населения в Новосибирской области за 2000-2019 года имеет вид [3]:

$$Y_t = 16,664e^{-0,016t}$$

Получаем $Y_{2020} = 11,99$, $Y_{2021} = 11,81$.

В 2020 г. Р-оценка составила 21,8%, в 2021 г. - 41,5%, т.е. на 21,8% и на 41,5% больше смертей на 1000 населения, чем ожидалось, в 2020 и в 2021 годах соответственно.

Кроме того, существует мультипликатор ковидной смертности (covid mortality multiplier – СММ). СММ основывается на предположении о том, что ковид влечет за собой дополнительные смерти, непосредственно не вызванные коронавирусом. Мультипликатор выглядит следующим образом (M_i – число умерших от всех причин в изучаемом промежутке времени, M_b – число умерших от всех причин в аналогичный период предшествующего года или средняя величина за несколько лет до изучаемого периода, M_c – число умерших от

COVID-19 в рассматриваемом периоде времени, $\Delta M_g = M_i - M_b$ – прирост или сокращение общего числа умерших) [4].

$$CMM = \frac{(M_i - M_b)}{M_c} = \frac{\Delta M_g}{M_c}$$

Чем выше полученное значение CMM, тем заболевание более смертельно.

Мультипликаторы ковидной смертности для Новосибирской области представлены на рисунке 1.

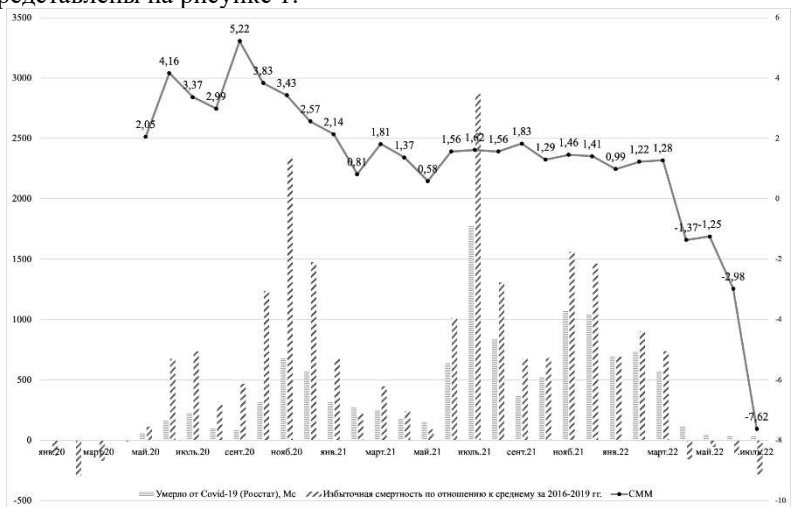


Рис. 1 - Числа умерших от COVID-19, избыточная смертность и коэффициент CMM по месяцам 2020-2022 гг. в Новосибирской области

В сентябре 2020 года зафиксировано наибольшее значение мультипликатора ковидной смертности в Новосибирской области – он составил 5,22. Это означает, что в среднем каждый заболевший COVID-19 на тот момент мог привести к смерти еще пяти человек.

Таким образом, существует несколько подходов к оценке избыточной смертности. Такой показатель, как избыточная смертность, может быть полезен при оценивании последствий не только пандемий, но и войн, природных катастроф и других чрезвычайных ситуаций.

Литература:

1. Глобальная избыточная смертность, связанная с пандемией COVID-19 [Электронный ресурс] // Всемирная организация здравоохранения. – 2022. – URL: <https://www.who.int/ru/news-room/questions-and-answers/item/global-excess-deaths-associated-with-the-COVID-19-pandemic> (дата обращения: 28.10.2024).
2. Джаттино Ч. Повышенная смертность во время пандемии коронавируса (COVID-19) [Электронный ресурс] // Our World in Data – 2020. – URL: <https://ourworldindata.org/excess-mortality-covid> (дата обращения: 23.10.2024).
3. Избыточная смертность в Новосибирской области [Электронный ресурс] // GOGOV – о главном в России без политики – 2023. – URL: <https://gogov.ru/excess-mortality/nso#data> (дата обращения: 29.10.2024).
4. Кашенов А.В. Избыточная смертность населения во время пандемии COVID-19 в регионах России [Электронный ресурс] // Вестник Алтайской академии экономики и права. – 2022. – № 9-3. – С. 352-360. – URL: <https://vaael.ru/ru/article/view?id=2481> (дата обращения: 20.10.2024).
5. Территориальный орган Федеральной службы государственной статистики по Новосибирской области.

ОПИСАНИЕ ПРОФИЛЯ РЕЛЬСА В ПРОЦЕССЕ ИЗНОСА С ПОМОЩЬЮ РЕГРЕССИОННЫХ СПЛАЙНОВ

А.А. Кривецкий, В.С. Карманов

**Новосибирский государственный технический университет,
г. Новосибирск, kriveczkij.2020@stud.nstu.ru**

Представлена методика построения кубических регрессионных сплайнов. Рассмотрены результаты численного эксперимента аппроксимации профиля изношенного рельса для различного количества интервалов. Приведена методика получения и обработки исходных данных для изношенного рельса.

The technique of cubic regression splines construction is presented. The results of numerical experiment of approximation of worn rail profile for different number of intervals are considered. The technique of obtaining and processing of initial data for the worn rail is given.

Введение

В процессе эксплуатации железнодорожный рельс подвержен износу [1]. Для применения технологий выравнивания профиля [2, 3], необходимо иметь простую, непрерывную модель описания поверхности. Для этого предлагается аппроксимировать профиль рельса кубическим регрессионным сплайном.

Регрессионный сплайн

Для аппроксимации данных будем использовать уравнение полиномиальной регрессии третьего порядка, линейное по параметрам:

$$y_i = f(x_i) = \theta_0 + \theta_1 x_i + \theta_2 x_i^2 + \theta_3 x_i^3 + \varepsilon_i, \quad i = 1..n, \quad (1)$$

где x_i – независимые переменные (регрессоры), а y_i – зависимые (значения отклика), $i = 1..n$.

Тогда оценки параметров найдутся по формуле [4]:

$$\hat{\theta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}, \quad (2)$$

$\theta = (\theta_0, \theta_1, \dots, \theta_m)'$ – вектор-столбец неизвестных параметров,

$\hat{\theta}$ – оценка вектора параметров,

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & x_n^3 \end{pmatrix} \text{ – матрица плана (наблюдений),}$$

\mathbf{X}' – транспонирование матрицы,

$\mathbf{Y} = (y_1, y_2, \dots, y_n)'$ – вектор значений (отклика).

После оценки параметров, можно построить значения откликов аппроксимирующего полинома следующим образом:

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\theta}.$$

Рассмотрим метод построения регрессионного кубического сплайна. Для построения необходимо разбить исходные данные x_i на интервалы $\Delta_k = [u_k, u_{k+1}]$, $k = 1..L$, где L – количество интервалов, u_k – границы интервалов. Они же (кроме точек u_1 и u_{L+1}) будут являться узлами сплайна, в этих точках будет обеспечиваться «стыковка» уравнений регрессии на соседних интервалах, до производной второго порядка. Для сопоставления уравнений регрессии в узлах, необходимо задать систему линейных ограничений:

$$\mathbf{R}\theta = \mathbf{r}. \quad (3)$$

В случае построения регрессионного сплайна, матрица плана – блочная и имеет следующий вид:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{X}_D \end{pmatrix},$$

где \mathbf{X}_k , $k = 1..L$ – матрица плана для построения регрессии (1) на интервале Δ_k , имеет следующий вид:

$$\mathbf{X}_k = \begin{pmatrix} 1 & x_{1,k} & x_{1,k}^2 & x_{1,k}^3 \\ 1 & x_{2,k} & x_{2,k}^2 & x_{2,k}^3 \\ \dots & \dots & \dots & \dots \\ 1 & x_{n,k} & x_{n,k}^2 & x_{n,k}^3 \end{pmatrix},$$

где $x_{i,k}$ – i -ое значение наблюдения из интервала Δ_k .

Рассмотрим равномерное разбиение выборки. Так как $\Delta_k = [u_k, u_{k+1}]$, для нахождения узлов сплайна воспользуемся формулой:

$$u_k = x_1 + (k-1)h, \quad k = 1 \dots L+1$$

где $h = \frac{x_n - x_1}{n}$ – длина интервала.

Матрица ограничений \mathbf{R} – блочная, имеет следующий вид:

$$\mathbf{R} = \begin{pmatrix} \mathbf{G}(u_2) & -\mathbf{G}(u_2) & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}(u_3) & -\mathbf{G}(u_3) & \dots & \mathbf{0} & \mathbf{0} \\ & \vdots & & \ddots & & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{G}(u_L) & -\mathbf{G}(u_L) \end{pmatrix},$$

где матрица $\mathbf{G}(u)$ будет выглядеть следующим образом:

$$\mathbf{G}(u) = \begin{pmatrix} 1 & u & u^2 & u^3 \\ 0 & 1 & 2u & 3u^2 \\ 0 & 0 & 2 & 6u \end{pmatrix}.$$

Вектор \mathbf{r} будет нулевым и оценка, минимизирующая сумму квадратов отклонений при ограничениях (3), вычисляется по формуле [4]:

$$\hat{\boldsymbol{\theta}}_{\mathbf{R}} = (\mathbf{E} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R})\hat{\boldsymbol{\theta}}, \quad (4)$$

где \mathbf{E} – единичная матрица.

Численный эксперимент

В качестве иллюстрации предлагаемого подхода используются данные из открытых источников [5], рис. 1. Была проведена оцифровка данного изображения при использовании методов библиотеки OpenCV. Разрешение изображения было предварительно увеличено.



Рис. 1 – Профиль рельса, подверженного износу

На рис. 2 представлен пример работы метода. Используется равномерное разбиение выборки на 8 интервалов. В качестве показателя ошибки рассматривается величина $e = Y - \hat{Y}$.

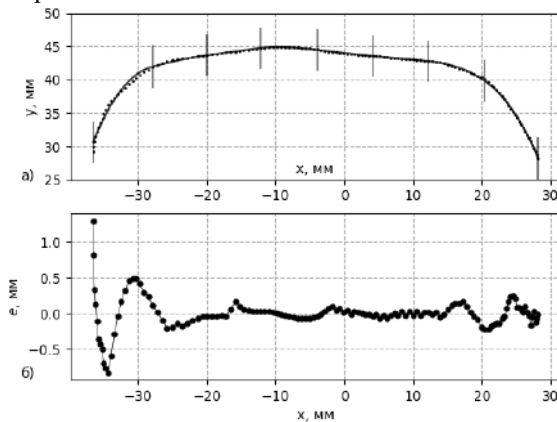


Рис. 2 – Аппроксимация поверхности рельса. 8 интервалов.
а) Регрессионный сплайн и исходные данные, б) Ошибка аппроксимации.

В качестве показателей точности аппроксимации, рассмотрим среднеквадратическое отклонение (RMSE – Root Mean Square Error) и модуль максимального отклонения (MaxAE – maximum absolute error), которые вычисляются по формулам:

$$RMSE = \sqrt{\frac{1}{n}(\mathbf{Y} - \hat{\mathbf{Y}})'(\mathbf{Y} - \hat{\mathbf{Y}})},$$

$$MaxAE = \max|\mathbf{Y} - \hat{\mathbf{Y}}|.$$

В таблице 1 приведено сравнение показателей точности при разном количестве интервалов.

Таблица 1 Сравнение показателей точности при разном количестве интервалов

Количество интервалов	RMSE, мм	MaxAE, мм
3	0,601	2,551
6	0,271	1,515
9	0,220	1,171
12	0,174	0,919
15	0,124	0,681
18	0,087	0,480

Как и ожидается, с увеличением числа интервалов, увеличивается точность аппроксимации. Например, для абсолютной точности не выше 1 мм рекомендуется использовать 12 интервалов. При определении числа интервалов, необходимо учитывать, что для составления регрессии по указанной методике на каждом интервале должно быть не менее 5 точек.

Заключение

Разработанный подход может найти широкое применение при обработке железнодорожных рельсов и в области строительства и анализа качества железных дорог.

Литература:

1. *Шур Е.А.* Перспективные требования российских железных дорог к рельсам / Е.А. Шур // Железнодорожный транспорт. – 2008. – № 2. – С. 41-45.
2. *Бунтаков Д.В.* Методы обработки материалов для восстановления изношенности рельсового пути [Электронный ресурс] // Вестник науки. – 2023. – Т. 4, №12 (69). – С. 1143-1152. – URL: <https://www.вестник-науки.рф/article/11996> (дата обращения: 08.11.2024).
3. *Иноземцев В.Е., Попов А.Ю., Кульков А.А., Корытов А.Ю.* Репрофилирование железнодорожных рельсов на рельсофрезерных станках // Транспортное машиностроение. – 2022. – №11 (11).
4. *Демиденко Е.З.* Линейная и нелинейная регрессии / Е.З. Демиденко. – М.: Финансы и статистика, 1981. – 302 с.

5. Распоряжение ОАО «РЖД» от 23.10.2014 № 2499р (с изм. от 05.10.2020) «Об утверждении и введении в действие инструкции «Дефекты рельсов. Классификация, каталог и параметры дефектных и острodefектных рельсов» // СПС КонсультантПлюс.

ОПТИМИЗАЦИЯ ПЛАНИРОВАНИЯ ПОСЕВОЧНЫХ РАБОТ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ ОПТИМИЗАЦИИ

Е.Е. Марчев, Д.М. Плескач, А.В. Фаддеев

**Заполярный государственный университет им. Н.М. Федоровского,
г. Норильск, mrc Chew228057@gmail.com**

В данной работе описывается разработка программного комплекса для оптимального распределения сельскохозяйственных культур на полях с использованием различных методов оптимизации. Цель проекта — автоматизировать процесс выбора культур с учётом характеристик полей и условий выращивания для увеличения урожайности и снижения затрат. Исследования проводятся на основе нескольких алгоритмов: полного перебора, жадного алгоритма, метода ветвей и границ, а также генетического алгоритма, что позволяет сравнить их эффективность в различных условиях.

This paper describes the development of a software system for the optimal distribution of agricultural crops across fields using various optimization methods. The goal of the project is to automate the crop selection process by considering field characteristics and growth conditions to increase yield and reduce costs. Research is conducted based on several algorithms: exhaustive search, greedy algorithm, branch and bound, and genetic algorithm, enabling a comparison of their efficiency under different conditions.

Цель работы

Целью исследования является создание программного комплекса для выбора оптимального распределения посевных культур по сельскохозяйственным полям. Программа ориентирована на применение в условиях, требующих быстрой и точной оценки параметров культур и полей для обеспечения максимальной прибыли при сохранении баланса использования ресурсов.

Введение

В современном сельском хозяйстве возрастают требования к эффективности использования сельскохозяйственных ресурсов, в том числе посевных площадей, водных и энергетических ресурсов, а также к максимизации экономической отдачи от посевных культур. Применение интеллектуальных систем управления, основанных на алгоритмах оптимизации, позволяет достичь высоких результатов за счёт более

точного распределения культур в соответствии с их характеристиками и параметрами полей.

В данной работе описывается разработка автоматизированного программного комплекса для управления распределением посевных культур с использованием различных алгоритмов оптимизации.

Постановка задачи

Программный комплекс решает задачу выбора культур для посева с учётом характеристик сельскохозяйственных полей, таких как:

- Площадь поля.
- Тип почвы.
- Доступность водных ресурсов.
- Ожидаемая прибыль от каждой культуры.

На вход программе подаются параметры каждого поля и возможные культуры, из которых необходимо выбрать наиболее подходящие. На основе целевой функции оценивается эффективность распределения культур. Вычисления производятся с использованием таких методов, как полный перебор, жадный алгоритм, метод ветвей и границ, и генетический алгоритм, что позволяет гибко учитывать размеры и сложность задач.

Описание алгоритмов и результаты

Были проведены сравнительные испытания алгоритмов, результаты которых показали, что каждый метод имеет свои преимущества. Полный перебор и метод ветвей и границ обеспечивают высокую точность, но требуют значительных временных затрат. Жадный алгоритм и генетический алгоритм показали хорошие результаты по скорости, но могут давать менее точные результаты при больших наборах данных.

Алгоритм	Точность (%)	Время выполнения (мс)
Полный перебор	100	9200
Жадный алгоритм	85	7300
Метод ветвей и границ	95	4500
Генетический алгоритм	92	4000

Заключение

Разработанная система позволяет проводить автоматизированный выбор посевных культур для конкретных полей с учётом множества

параметров, что способствует повышению прибыли и оптимизации использования сельскохозяйственных ресурсов. На основании представленных результатов можно сделать вывод, что в зависимости от задач рекомендуется использовать различные алгоритмы: полный перебор и метод ветвей и границ – для задач с высоким приоритетом точности, генетический и жадный алгоритмы – для задач, требующих быстрого результата.

Литература:

1. Ритчи Р.К., Чандлер Д. Методы оптимизации: от жадных до эволюционных. – СПб.: Питер, 2017.
2. Кормен Т.Х., Лейзерсон Ч.Э., Ривест Р.Л., Штайн К. Алгоритмы: построение и анализ. – М.: Вильямс, 2020.
3. Стендер М. Эволюционные алгоритмы и их применение в оптимизационных задачах. – М.: Мир, 1991.

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТЕЙ ПРИМЕНЕНИЯ СТАТИСТИЧЕСКИХ МЕТОДОВ ДЛЯ КОРПУСОВ ФИЛОСОФСКИХ И ХУДОЖЕСТВЕННЫХ ТЕКСТОВ

И.Д. Мусаткин

Новосибирский государственный технический университет,

г. Новосибирск, musatkin.2021@stud.nstu.ru

Научный руководитель: Тимофеева А.Ю., к.э.н., доцент

В данной статье демонстрируется, полезность частотного анализа и NLP для философских и литературных исследований в определении динамики ключевых тем и терминологического анализа авторов

This article demonstrates the usefulness of frequency analysis and NLP for philosophical and literary research in determining the dynamics of key topics and terminological analysis of authors

Статистический анализ, применяемый к текстовым данным, машинное обучение и NLP уже широко используются в задачах машинного перевода, создания чат-ботов, для решения проблем бизнеса и медицины. С помощью этих методов также проводятся исследования философии и литературы [1][2], но нереализованных возможностей и нераскрытых сфер (текстов, авторов) для исследования в этих областях – всё ещё очень много.

В данной работе мы сформировали для анализа корпус из философских и художественных текстов, и попытались выполнить ряд

задач: 1) проверить наличие взаимосвязей между авторами путём терминологического сравнения; 2) выявить цикличные темы в произведении путём анализа встречаемости соответствующих терминов на протяжении текста; 3) выявить разницу в лексиконе философского и прозаического жанров на примере одного автора.

Для выполнения этих задач был реализован многоэтапный процесс предобработки текста методами обработки естественного языка (NLP). Затем применены частотные методы анализа биграмм и униграмм для корпуса художественных и философских текстов и на их основе реализованы различные методы визуализации для поставленных задач.

Первый корпус текстов для выявления терминологических связей был сформирован из основных произведений 5 философов различных времён и направленностей: Иммануила Канта, Фридриха Ницше, Альбера Камю, Жюль Делёза и Мераба Мамардашвили.

Эмпирическим материалом для второго корпуса было выбрано одно из главных произведений 20 века – семитомный роман Марселя Пруста «В поисках утраченного времени». Текст романа был взят на языке оригинала, ввиду отсутствия на русском языке полного перевода, сделанного одним автором (не содержащим терминологических расхождений).

Для того, чтобы произвести частотный анализ, необходимо привести текст в подходящую для обработки форму – для этого были программно реализованы: токенизатор текста, удаление служебных символов, словарь стоп-слов, лемматизация и частеречная разметка. Основу для словаря стоп-слов взяли в библиотеке для обработки естественного языка на *python* – *nltk*, и добавили необходимые для специфики работы с художественными текстами слова. Лемматизацию произвели с помощью *core_news_sm* – модели для обработки естественного языка из библиотеки *SpaCy*. В каждой из библиотек существуют модификации как для русского (первый корпус текстов), так и для французского (для романа Пруста) языков.

Анализ биграмм оказался малоэффективен на материале художественного текста, причиной этому можно выделить ряд причин:

- наибольшие показатели лексической уникальности среди всех жанров текстов
- большое количество стилистических замещений, синонимов и недосказанности (из-за чего задача выделения сквозных тем становится более трудной)
- невозможность сформировать список терминов для оценки успешности проверки биграмм статистическими гипотезами о

независимости (какой можно составить для анализа текстов предметной области с определёнными границами и общепринятыми терминами).

Тем не менее, количественный анализ биграмм можно использовать для визуализации относительной частоты биграмм-терминов в текстах двух авторов. Сравнив тексты Фридриха Ницше и Жюль Делёза, можно увидеть, что последний заимствует и развивает в своих текстах такие термины Ницше, как «вечное возвращение», «воля к власти», «аскетический идеал».

Большую эффективность показал количественный анализ униграмм – при условии его конкретизации на терминах-существительных. Для этого мы использовали *POS-tagging* – с помощью упомянутых выше модификаций модели *core_news_sm*. Как правило, именно существительные являются терминами, так что такое ограничение позволит устранить из анализа множество служебных конструкций, не отсеянных на этапе обработки стоп-словарём.

Мы получили частотность терминов для текстов Альбера Камю - автора, писавшего как романы, так и философские трактаты. Разделив его тексты на 2 корпуса по этому критерию, посмотрим на разницу униграмм, сравнив также с корпусом Фридриха Ницше, повлиявшего на Камю.

Таблица 1 – фрагмент частотного анализа лексики жанров, значения – относительные частоты терминов $\times 1000$

Слово	Ницше	Проза Камю	Философия Камю
дух	6.6	1.4	8.1
конец	5.8	9.7	10.7
власть	5.8	2.0	6.2
цель	5.5	1.0	5.7
истина	5.4	1.5	3.5
отношение	5.3	1.2	4.0

Можно увидеть, что философские произведения Ницше и Камю сходятся в таких общих для философии темах, как «истина», «дух», «цель», которые не так часты в художественных текстах последнего.

Визуализировав количественный анализ униграмм для корпусов разных авторов, можно увидеть сходство и различие в ключевых терминах:

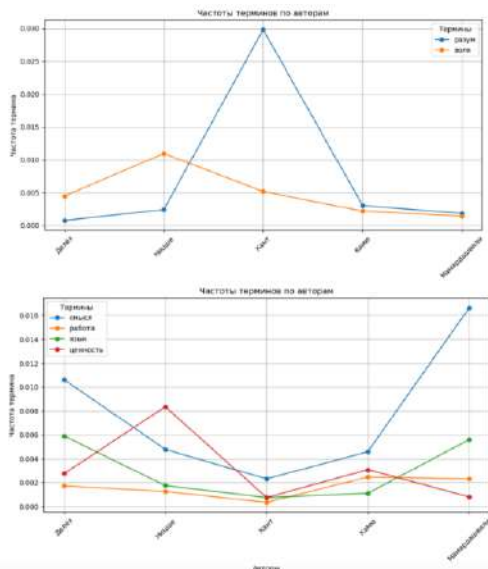


Рис.1 – относительная частота терминов по авторам

На первом графике на рисунке 1 можно явно выделить «философов воли» - Ницше и Делёза и «философа разума» - Канта. На втором графике - увидеть, что такие термины, как «язык», «ценность» - получают большее распространение в философии 19-20 века. У Канта, принадлежащего по преимуществу веку 18, наименьшая относительная частота для каждого из рассматриваемых терминов.

Ещё одна возможность применения анализа униграмм – визуализация частот слов на протяжении всего текста Пруста.

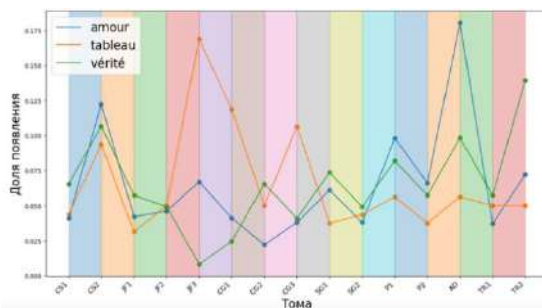


Рис.2 – появление термина на протяжении романа

На рисунке визуализация построена по относительным частотам терминов в конкретных томах к их общему количеству. Можно заметить, что темы любви (amour), истины (vérité) и живописи (tableau - картина) – постоянны в романе. Помимо этого, на графике видно и том романа, в котором больше всего рассуждений на определённую тему. К примеру, термин «картина» преобладает в 5 томе, в котором в действительности много фигурирует живопись.

Как показали результаты исследования, даже с помощью несложных в реализации алгоритмов, можно получить полезные сведения о текстах, способные упростить работу исследователя в анализируемых предметных областях, например, выявить места с наибольшей концентрацией термина. Ещё большая область для анализа открывается с применением методов машинного обучения для тематического анализа, программная основа для реализации которых была заложена в рамках данного исследования.

Литература:

1. Moretti F. Graphs, Maps, Trees. Abstract Models for a Literary History – Verso, 2005. – 119 p.
2. Макулин А.В. Интеллектуальные системы в гуманитарной сфере и цифровая философия // Вестник Северного (Арктического) федерального университета. Серия: Гуманитарные и социальные науки. 2016. – №2.

АНАЛИЗ ПРОФИЛЕЙ ВОЛН НА СВАРОЧНЫХ ШВАХ ПРИ СВАРКЕ ВЗРЫВОМ

Н.В. Обидин

Новосибирский государственный технический университет,

г. Новосибирск, obidinnikolaj1@gmail.com

Научный руководитель: Исаева Е. В., к.т.н.

В данной работе предложено решение задачи анализа профилей волн на сварных швах, полученных при сварке взрывом. Цель исследования — разработка автоматизированного метода для вычисления параметров волн (средней длины и амплитуды) на основе изображений швов, что повысит точность и эффективность оценки их качества. Предложенный подход включает обработку изображений, вычисление параметров, визуализацию и массовую обработку данных для полной автоматизации анализа.

This paper proposes a solution for analyzing wave profiles on welds produced by explosion welding. The goal is to develop an automated method to calculate wave parameters (average wavelength and amplitude) from weld images, enhancing the accuracy and efficiency of quality assessment. The approach involves image processing, parameter calculation, result visualization, and batch data processing to fully automate the analysis.

Введение

Анализ качества сварных швов важен для обеспечения надёжности соединений. При сварке взрывом на шве формируются волны, характеристики которых, такие как длина и амплитуда, помогают оценить качество соединения. В работе использованы изображения моделирования сварки двух металлов взрывом (Рисунок 1), а также изображения реального эксперимента сварки этих металлов взрывом (Рисунок 2).

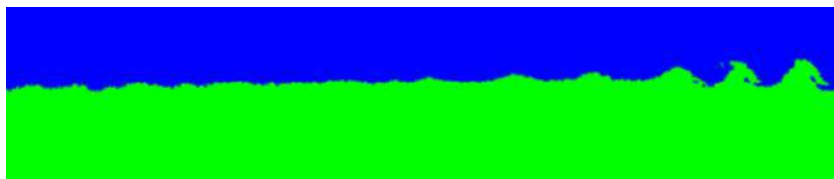


Рис. 1. - Изображение моделирования процесса сварки двух металлов взрывом.

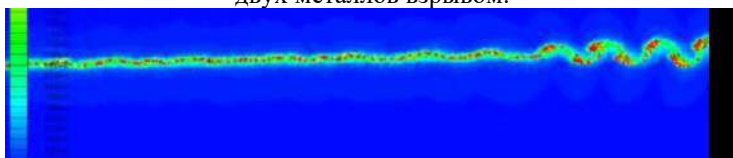


Рис. 2. - Изображение реального эксперимента сварки двух металлов взрывом.

Одной из проблем при анализе данных изображений является шум, искажающий результаты. Для его уменьшения используется метод Гауссова размытия.

Гауссово размытие — это метод обработки изображений, основанный на применении Гауссова фильтра. Этот фильтр действует на изображение путем размытия пикселей, что помогает смягчить резкие контуры и шум. Функция фильтра Гаусса [1] описывается следующим образом:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}}, \quad (1)$$

где $G(x, y)$ — значение фильтра в точке (x, y) , σ — стандартное отклонение, определяющее степень размытия.

Применение данного фильтра позволяет устранить мелкие шумы и подготовить изображение к дальнейшему анализу, сохраняя при этом ключевые особенности профиля.

После предобработки изображения выполняется извлечение 1D профиля волны, что позволяет упростить анализ и сосредоточиться на ключевых характеристиках шва. Для этого вычисляется усредненное значение интенсивности по строкам изображения:

$$P(x) = \frac{1}{N} \sum_{i=1}^N I(i, x)$$

где $P(x)$ — одномерный профиль интенсивности, $I(i, x)$ — интенсивность пикселя на строке i и колонке x , N — количество строк изображения.

Полученный профиль представляет собой последовательность значений, отражающих изменение интенсивности вдоль шва.

Для анализа профиля волн необходимо определить их параметры, такие как длина волны и амплитуда, путем поиска локальных экстремумов — пиков и впадин.

Пик — точка, где первая производная равна нулю, при этом производная перед точкой положительна, а после — отрицательна (максимум).

Впадина — точка, где первая производная также равна нулю, но производная перед точкой отрицательна, а после — положительна (минимум).

Знак первой производной определяет тип экстремума. Средняя длина волны вычисляется как расстояние между последовательными пиками.

$$\lambda = \frac{1}{n-1} \sum_{i=1}^{n-1} (p_{i+1} - p_i),$$

где n — количество найденных пиков, p_i — позиции пиков на профиле.



Рис. 3. - Профиль волны с пиками и впадинами для изображения моделирования сварки двух металлов взрывом.

Средняя амплитуда определяется как разница между значениями профиля в пиках и впадинах:

$$A = \frac{1}{m} \sum_{i=1}^m (P(p_i) - P(t_i))$$

где m — количество пар "пик-впадина", $P(p_i)$ и $P(t_i)$ — значения профиля на пиках и впадинах соответственно.



Рис. 4. - Профиль волны с пиками и впадинами для изображения реального эксперимента сварки двух металлов взрывом.

В таблице (Таблица 1) представлены результаты средних длин волн и амплитуд, вычисленных для изображений, полученных в процессе моделирования и реального эксперимента сварки двух металлов взрывом.:

**Таблица 1. Сравнительный анализ средних длин волн
и амплитуд для изображений**

Изображение	Средняя длина волны	Средняя амплитуда волны
Рис. 3.	11.0161290322580	1.626666666666666
Рис. 4.	8.49367088607595	2.33299578059071

Заключение

В данной работе предложен метод автоматизированного анализа профилей волн на сварных швах, полученных при сварке взрывом. Разработан алгоритм, включающий Гауссово размытие для уменьшения шума, извлечение 1D профиля волны и вычисление ключевых параметров, таких как средняя длина волны и амплитуда. Проведены тестирования метода и оценка его эффективности для автоматизированной обработки изображений и оценки качества сварных соединений.

Литература:

1. *Фельдман А.В.* Цифровая обработка сигналов. – М.: Радио и связь, 2003. – 400 с.
2. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. – М.: Мир, 2014. – 800 с.
3. *Цисарж В.В., Марусик Р.И.* Математические методы компьютерной графики: учеб. пособие. – К.: Факт, 2004. – 466 с.

СТАТИСТИЧЕСКИЙ АНАЛИЗ ПРЕСТУПНОСТИ В РОССИЙСКОЙ ФЕДЕРАЦИИ ЗА 2014-2023 ГГ.

В.М. Рубан

**Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, valeriiaaruban@yandex.ru**

Научный руководитель: Шмарихина Е.С., к.э.н., доцент

Анализ преступности позволяет выявить тенденции, определить факторы, влияющие на преступную активность, и сформировать рекомендации для повышения эффективности правопорядка. В данной работе рассматриваются ключевые аспекты преступности за последние десять лет. Эти данные позволяют не только оценить текущее состояние правопорядка, но и спрогнозировать возможные изменения в уровне преступности, что

способствует формированию обоснованной стратегии борьбы с преступностью.

Crime analysis makes it possible to identify trends, identify factors influencing criminal activity, and make recommendations to improve the effectiveness of law and order. This paper examines the key aspects of crime over the past ten years. These data make it possible not only to assess the current state of law and order, but also to predict possible changes in the crime rate, which contributes to the formation of a sound crime control strategy.

Преступность является важной социальной проблемой, оказывающей значительное влияние на стабильность и безопасность общества. Под преступностью понимается совокупность противоправных деяний, совершаемых на определенной территории за определенный период, которые подлежат наказанию в соответствии с законом. Анализ уровня преступности и его изменений позволяет не только понять текущую ситуацию, но и оценить эффективность деятельности правоохранительных органов и государственных мер профилактики правонарушений.

Преступность включает в себя не только само нарушение закона, но и широкий спектр социальных, экономических и психологических факторов, которые провоцируют людей на противоправные действия. Уровень преступности зависит от множества аспектов: уровня жизни населения, уровня безработицы, доступности образования, культурных и социальных норм, а также качества работы правоохранительных органов [1].

Одним из ключевых аспектов анализа преступности является динамика количества зарегистрированных преступлений и количества предварительно расследованных преступлений (рисунок 1) [2].

По данным МВД, с 2014 по 2023 годы в Российской Федерации наблюдается устойчивое снижение как зарегистрированных, так и предварительно расследованных преступлений. После пика в 2015 году (2,39 млн зарегистрированных преступлений и 1,25 млн расследованных) к 2023 году показатели снизились до 1,95 млн (-20,5%) и 0,99 млн (-18,5%), соответственно. Замедление темпов снижения в последние годы указывает на стабилизацию уровня преступности.

Увеличение разрыва между числом зарегистрированных и расследованных преступлений до 2020 года (+4,3%) сменилось тенденцией к его сокращению (-5,6%) в последние годы, что свидетельствует о повышении раскрываемости и эффективности следственных мероприятий.

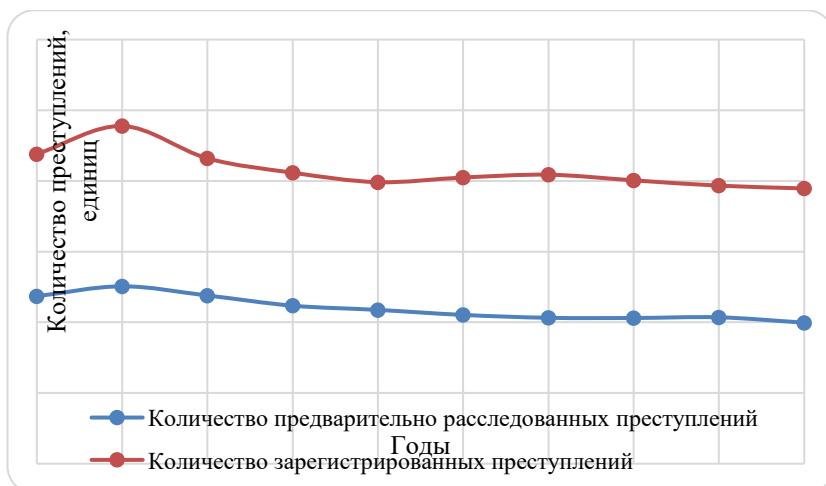


Рис. 1 – Динамика количества зарегистрированных и предварительно расследованных преступлений в России за 2014-2023 гг.

Структура преступности в России по данным 2023 года не имеет однородности (таблица 1), однако есть явно выраженные направления [3].

Таблица 1 – Структура зарегистрированных преступлений в России за 2014 и 2023 годы, в процентах

Вид преступления	2014	2023
Убийство	0,54	0,38
Умышленное причинение вреда здоровью	1,50	0,86
Изнасилование	0,19	0,16
Кража	41,49	29,96
Грабеж	3,55	1,14
Разбой	0,65	0,17
Вымогательство	0,21	0,46
Хулиганство	0,18	0,08
Экономической направленности	4,92	5,41
Мошенничество	7,27	22,27
Прочие*	39,49	39,11

* Данные взяты с сайта МВД России: <https://мвд.рф/>

Высоко распространены преступления, связанные с хищением: кража и мошенничество. В 2023 году по сравнению с 2014 годом структура преступности существенно изменилась ($I_R = 0,172$): сокращение насильственных и имущественных преступлений сопровождается резким ростом экономических преступлений, особенно мошенничества и вымогательства.

Для выявления закономерности в колебаниях уровня преступности в зависимости от времен года проведен анализ сезонности количества зарегистрированных преступлений за 2017-2023 годы (рисунки 2,3) [2].



Рис. 2 – Количество зарегистрированных преступлений в России по месяцам за 2017-2023 гг.

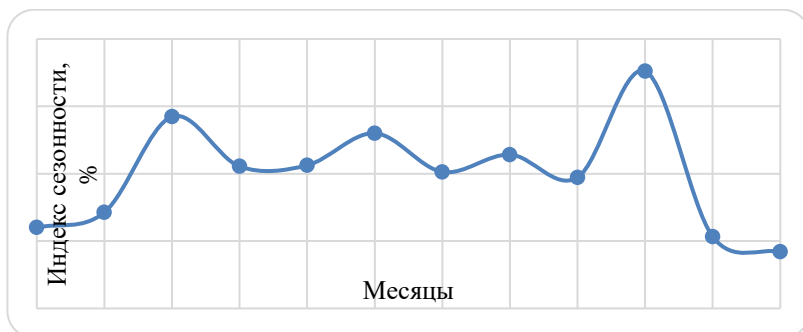


Рис. 3 – Показатели индекса сезонности преступности в России по месяцам

Наивысшие показатели уровня преступности наблюдаются в марте (108,5%) и октябре (115,2%), что может быть связано с сезонными изменениями в поведении людей, началом весенних активностей или увеличением социальных конфликтов осенью. Наиболее низкие

показатели преступности в ноябре (90,7%), декабре (88,4%) и январе (92,0%), когда происходит усиление контроля со стороны правоохранительных органов, а также в связи с климатическими условиями.

Анализ уровня преступности в России за последнее 10-летие показывает заметные изменения в динамике, структуре и сезонности преступлений. Для борьбы с преступностью необходимо учитывать структурные изменения и сезонные колебания, что позволит более эффективно распределять ресурсы и разрабатывать стратегии, направленные на повышение уровня безопасности в стране.

Литература:

1. Санташов А.Л., Петрова И.А., Герасимова Е.В. Анализ показателей преступности в современной России // Журнал прикладных исследований. – 2020. – № 4. – С. 65–73.

2. Единая межведомственная информационно-статистическая система (ЕМИСС) [Электронный ресурс]. – URL: <https://www.fedstat.ru/> (дата обращения: 01.11.2024).

3. Министерство внутренних дел Российской Федерации [Электронный ресурс]. – URL: <https://мвд.рф/> (дата обращения: 01.11.2024).

ИССЛЕДОВАНИЕ РОБАСТНЫХ МОДИФИКАЦИЙ ЧАСТНОГО КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ

Д.А. Рыбачёк

**Новосибирский государственный технический университет,
г. Новосибирск, darie.r@bk.ru**

Научный руководитель: Тимофеева А.Ю., к.э.н.

В данной работе рассматриваются модификации коэффициента частной корреляции. Устойчивость модификаций проверяется на модельном примере. Модель строится таким образом, что парные корреляции между признаками практически нулевые, но коэффициент частной корреляции указывает на наличие значимой связи.

This paper examines modifications of the partial correlation coefficient. The stability of modifications is checked using a model example. The model is generated in such a way that the paired correlations between the features are practically zero, but the coefficient of partial correlation indicates the presence of a significant correlation.

Целью данной работы является исследование устойчивости робастных модификаций частного коэффициента корреляции.

Коэффициент частной корреляции показывает степень статистической связи двух переменных относительно друг друга после устранения влияния одной или нескольких переменных. Он может быть рассчитан на основе парных коэффициентов корреляции: коэффициента корреляции Пирсона или коэффициента ранговой корреляции Спирмена (являющимся более устойчивым к выбросам).

В предыдущей работе [2] для расчёта робастного коэффициента частной корреляции воспользовались критерием LOOT (leave-one-out test, критерий поэлементного удаления). Процедура поэлементного удаления повторяется n раз для набора данных, состоящего из n точек: из данных выбрасывается одна точка и считается коэффициент корреляции без удалённой точки. Таким образом, в результате работы критерия имеется n коэффициентов корреляции. На основе средневзвешенного значения полученных коэффициентов рассчитывается робастный коэффициент корреляции.

При помощи данной процедуры можно также подсчитать робастный коэффициент корреляции Пирсона и уже на основе его получить частный коэффициент корреляции при исключении всех остальных переменных по следующей матричной формуле:

$$R_{ij \setminus vars} = \frac{-R_{ij}}{\sqrt{R_{ii}R_{jj}}}, \quad (1)$$

где R_{ij} – алгебраическое дополнение элемента r_{ij} корреляционной матрицы R , равное $R_{ij} = (-1)^{i+j}M_{ij}$, где M_{ij} – минор элемента r_{ij} матрицы робастных коэффициентов корреляции Пирсона.

Другими стандартными методами получения робастных коэффициентов корреляции являются методы усечения. В этом исследовании усечённый коэффициент корреляции получается путём исключения части наиболее удалённых от основной массы наблюдений [1]. При задаваемой доле отсекаемых наблюдений α в упорядоченном ряду исключаются $\alpha/2$ минимальных и $\alpha/2$ максимальных наблюдений. По полученной подвыборке считается частный коэффициент корреляции.

В данной работе проводится исследование четырёх модификаций частного коэффициента корреляции:

- при помощи LOOT на основе парных коэффициентов корреляции Пирсона;

- при помощи LOOT на основе коэффициентов ранговой корреляции Спирмена;
- рассчитанный по формуле (1) через робастные коэффициенты корреляции Пирсона (получаемые при помощи критерия LOOT);
- усечённый коэффициент корреляции (с долей отсекаемых наблюдений 5%).

Исследование проводилось на модельных данных, генерируемых со следующими условиями: x и y , z и y – с нулевой парной корреляцией, но при исключении z частная корреляция между x и y является значимой. Парная корреляция между x и z близка к 1. Генерация модели проводилась следующим образом:

$$z = x + \varepsilon_z,$$

$$\text{где } x \sim N(0,1), D(\varepsilon_z) = \frac{1}{cor_{xz}^2} - 1, \varepsilon_z \sim N(0, D(\varepsilon_z)),$$

$$y = z + bx + \varepsilon_y,$$

$$\text{где } D(\varepsilon_y) = \frac{(b+1)^2}{cor_{xy}^2} - \left(b^2 + 2b + \frac{1}{cor_{xz}^2} \right), \varepsilon_y \sim N(0, D(\varepsilon_y)).$$

$$\text{Тогда } cor_{zy} = \frac{(b+1 + D(\varepsilon_z))}{\sqrt{(1 + D(\varepsilon_y)) \cdot ((b+1)^2 + D(\varepsilon_z) + D(\varepsilon_y))}}.$$

Чтобы обеспечить необходимые условия модельных экспериментов задавались следующие значения: $b = -0.98755$, коэффициенты корреляции $cor_{xz} = 0.98867$, $cor_{xy} = 0.0337$. Тогда $cor_{yz} = 0.0911$, а истинное значение частного коэффициента корреляции между x и y при исключении z получается $cor_{xy|z} = -0.3769$.

В полученную модель добавляются выбросы – случайная точка в x заменяется целым числом из промежутка от 4 до 10. Исследования проводились на уровнях выбросов равными 1%, 5% и 10%. Моделируется выборки объёмом 100. Модельные эксперименты повторяются 100 раз.

В процессе вычислительных операций было обнаружено, что в некоторых случаях частный коэффициент корреляции, рассчитанный по матричной формуле на основе робастного коэффициента корреляции Пирсона, оказывается больше 1, что противоречит определению коэффициента корреляции. Данная ошибка возникает из-за того, что получаемая матрица робастных коэффициентов корреляции Пирсона

оказывается несогласованной. Поэтому в данном исследовании подобная модификация признана неустойчивой и в дальнейшем подлежит доработке.

В таблице 1 приведены средние и среднеквадратичные (в скобках) значения абсолютных отклонений между r_{actual} и r_{Rob} в зависимости от доли выбросов, где r_{actual} – коэффициент частной корреляции модели без выбросов, рассчитываемый для каждой модификации, r_{Rob} – полученный робастный коэффициент корреляции.

Таблица 1 – Средние и среднеквадратичные значения абсолютных отклонений r_{actual} и r_{Rob}

Доля выбросов	Роб. частный на основе Пирсона	Роб. частный на основе Спирмена	Усечённый частный (5%)
1%	0.0061 (0.0076)	0.0091 (0.0145)	0.0097 (0.0097)
5%	0.2129 (0.1204)	0.1228 (0.0876)	0.1998 (0.115)
10%	0.2188 (0.1228)	0.1322 (0.1016)	0.2057 (0.1244)

Робастный частный коэффициент корреляции на основе Пирсона показывает наименьшее абсолютное отклонение при доле выбросов 1%, но с ростом количества выбросов значительно уступает другим модификациям. Робастный частный коэффициент корреляции на основе Спирмена проявляет себя лучше других модификаций при большой доле выбросов. Усечённый коэффициент оказывается менее устойчивым при доле выбросов 1%, но при большей доле показывает меньшее абсолютное отклонение по сравнению с робастными оценками на основе коэффициента корреляции Пирсона и значительно уступает робастным оценкам на основе коэффициента корреляции Спирмена. Это связано с тем, что в данном модельном примере все выбросы являются числами, большими, чем максимальная точка модели без выбросов и при работе усечённого коэффициента не все из них удаляются (так как при расчёте коэффициента из данных исключаются 2 минимальных и 3 максимальных наблюдения).

В данной работе проведено исследование устойчивости робастных модификаций частного коэффициента корреляции в зависимости от доли выбросов. При доле выбросов 1% наиболее устойчивым оказался робастный аналог коэффициента частной корреляции на основе коэффициента Пирсона, но с ростом количества выбросов лучше всего проявляет себя робастный частный коэффициент на основе ранговой корреляции Спирмена. Частный коэффициент корреляции, рассчитываемый на основе робастных коэффициентов корреляции Пирсона, был признан неустойчивым.

Литература:

1. Выявление устойчивых связей между показателями коннективности ЭЭГ и компонентами интеллекта / Т.В. Авдеевко, А.Ю. Тимофеева, М.Ш. Муртазина, О.М. Разумникова // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. – 2022. – Т. 11, №4. – С. 19-36.
2. Рыбачёк Д.А. Выявление робастных частных взаимосвязей по данным ЭЭГ / Д.А. Рыбачёк; науч. рук. А.Ю. Тимофеева. Наука. Технологии. Инновации: XVII Всероссийская научная конференция молодых ученых (г. Новосибирск, 04-08 декабря 2023 г.) // Сборник научных трудов в 11 ч. / Под ред. Соколовой Д.О. – Новосибирск: Изд-во НГТУ, 2024. – Ч. 2. – С. 119-122.

СИСТЕМА РЕЗЕРВИРОВАНИЯ СВОБОДНЫХ АУДИТОРИЙ

А.Н. Салимзянов, В.Г. Кобылянский

**Новосибирский государственный технический университет,
г. Новосибирск, asalimzanov14@gmail.com**

Научный руководитель: Кобылянский В.Г., к.т.н., доцент

Одной из проблем организации учебного процесса является поиск свободных аудиторий для проведения консультаций и других мероприятий, не предусмотренных расписанием занятий. В данной статье описаны результаты разработки веб-приложения, предназначенного для поиска и резервирования свободных аудиторий, относящихся к компетенции факультета.

One of the problems of organizing the educational process is the search for free classrooms for consultations and other activities that are not provided for in the class schedule. This article describes the results of developing a web application designed for searching and reserving available classrooms managed by the faculty.

Проведение внеплановых консультаций и других аудиторных мероприятий является неотъемлемой частью образовательного процесса. Для проведения консультаций обычно необходимы аудитории, оборудованные определенными техническими и программными средствами в соответствии с требованиями изучаемой дисциплины. При большом числе студенческих групп, изучающих дисциплину, проблема поиска и резервирования свободных аудиторий, относящихся к компетенции факультета, является актуальной и требует от преподавателя существенных затрат времени.

В данный момент для организации консультации преподаватель должен:

- провести анализ расписания всех студенческих групп, изучающих дисциплину, и найти день и время, свободные от занятий;
- найти в своем расписании день и время для проведения консультаций;
- лично получить от ответственного лица факультета сводное расписание занятости по всем аудиториям в бумажном виде и найти свободную аудиторию, удовлетворяющую результатам поисков, после чего ответственное лицо должно сделать отметку в сводном расписании о занятости аудитории на требуемый период времени.

Этот подход к организации консультаций имеет ряд существенных недостатков:

- сложность ручного согласования расписаний занятости аудиторий с расписанием преподавателей и студентов;
- необходимость личного присутствия преподавателя в университете;
- риск возникновения ошибок из-за человеческого фактора.

В настоящее время имеются многофункциональные коммерческие продукты, позволяющие решить эти проблемы, например TouchONE и SimplyBook [1,2]. Однако все они имеют избыточную функциональность и не учитывают особенности составления расписаний в разных учебных заведениях (нумерация недель, периодичность занятий, порядок резервирования аудиторий и т.д.).

Для решения проблемы поиска свободных аудиторий и их резервирования было разработано веб-приложение, реализующее следующий алгоритм:

- администратор заносит сводное расписание по всем аудиториям в базу данных;
- преподаватель подает заявку на резервирование аудиторий;
- администратор обрабатывает заявку, отклоняя или принимая её.

При работе над клиентской частью приложения были использованы следующие свободно распространяемые инструменты:

- React, который позволяет разрабатывать одностраничные веб-приложения;
- MobX для управления состоянием приложения;
- Axios для формирования и отправки HTTP-запросов на сервер;
- Bootstrap для построения адаптивного интерфейса.

Серверная часть разработана на Node.js в связке с фреймворком Express для реализации API. В качестве СУБД используется PostgreSQL. Для взаимодействия с базой данных применяется ORM Sequelize. Управление сессиями и правами доступа пользователей реализовано с использованием технологии JSON Web Token (JWT).

Разработанное приложение опубликовано в облачной платформе НГТУ и состоит из двух частей: основного окна и панели администратора. Доступ к функционалу каждой из этих частей ограничен соответствующими правами. На рисунке 1 представлена модель взаимодействия клиента и сервера.



Рис.1 – Модель клиент-серверного взаимодействия приложения

На рисунке 2 представлена панель администратора, с помощью которой он может вносить изменения в базу данных и управлять статусами заявок.

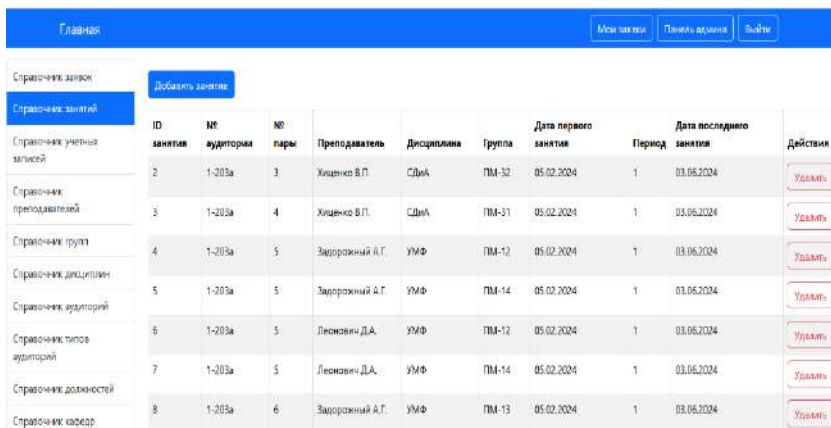


Рис. 2 – Панель администратора

Основное окно приложения представлено на рисунке 3. С помощью основного окна преподаватели могут переключаться между различными форматами отображения расписания, за счет чего обеспечивается удобство поиска свободных аудиторий. В этом же окне преподаватели могут подавать заявки на резервирование аудиторий. Приложение также позволяет резервировать аудитории с определенным периодом.

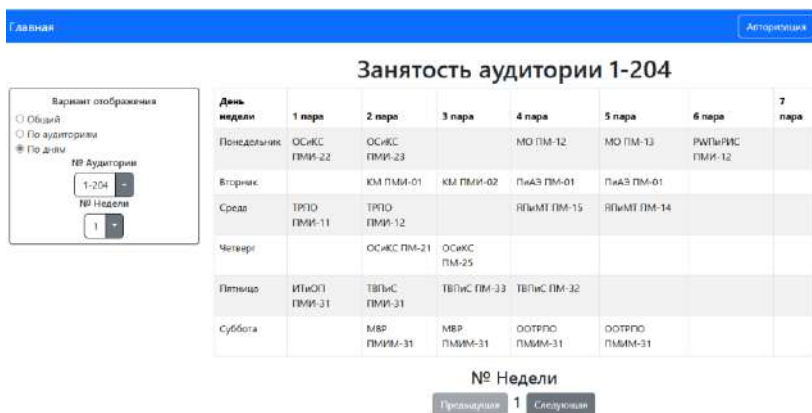


Рис. 3 – Основное окно приложения

Новый порядок действий для преподавателя, реализованный в приложении, выглядит следующим образом:

- авторизация в системе;

- поиск даты и времени в электронном сводном расписании занятости аудиторий;
- ввод данных по проводимому мероприятию с помощью всплывающей формы с использованием справочников (фамилия преподавателя, название дисциплины, периодичность и т.д.);
- отправка заявки.

В дальнейшем планируется развитие системы путем добавления возможности согласования расписания занятости аудиторий с занятостью преподавателей и студентов, а также получение расписания напрямую из информационной системы университета.

Литература:

1. Приложение онлайн-записи для университетов SimplyBook [Электронный ресурс] // simplybook.me. – URL: <https://simplybook.me/ru/appointment-scheduling-software-for-educational-services/appointment-scheduling-app-for-universities> (дата обращения: 04.11.2024).
2. Система резервирования TouchONE // touchone.ru. – URL: <https://touchone.ru/> (дата обращения: 04.11.2024).

СПОСОБ ИЗУЧЕНИЯ НЕРАВНОМЕРНОГО РАСПРЕДЕЛЕНИЯ ДОХОДОВ С ПРИМЕНЕНИЕМ МЕТОДА СОРТИРОВКИ ПО ПРОЦЕНТИЛЯМ

Т.В. Семенов, А.В. Фаддеенков

**Заполярный государственный университет им. Н.М. Федоровского,
г. Норильск, semtimsecondadd@gmail.com**

В статье предлагается новый подход к изучению неравномерного распределения доходов, основанный на методе сортировки по процентиллям. Этот метод позволяет выявить и детально проанализировать различия между группами доходов с высокой степенью точности. Применение данного подхода даёт возможность более точно анализировать распределение доходов, в частности, в самых низких и самых высоких группах, что актуально для оценки уровня неравенства в обществе.

The article proposes a new approach to the study of uneven income distribution based on the percentile sorting method. This method allows you to identify and analyze in detail the differences between income groups with a high degree of accuracy. The application of this approach makes it possible to more accurately analyze the distribution of income, in particular, in the lowest and highest groups, which is relevant for assessing the level of inequality in society.

Неравномерное распределение доходов является одним из ключевых вопросов экономической теории и политики, поскольку оно оказывает влияние на социальное благополучие и экономическую стабильность общества. Анализ распределения доходов позволяет понять, насколько велико расслоение общества, как формируется структура доходов различных групп и на какой стадии находится экономика с точки зрения социального равенства [1][4].

Среди методов изучения неравномерности распределения доходов особое внимание уделяется подходам, основанным на процентилях, что позволяет выявлять значительные отклонения от равномерности распределения. В данной статье рассматривается применение метода сортировки доходов по процентилям для изучения неравномерного распределения, а также выделяются основные преимущества и ограничения этого подхода.

Метод процентилей позволяет разбить совокупность доходов на определённые части, каждая из которых содержит равное количество наблюдений. Например, при использовании квартилей или децилей доходы можно распределить по четырём или десяти равным частям соответственно. Этот подход позволяет выявить, какая часть дохода концентрируется в верхнем или нижнем диапазоне распределения. В частности, анализ процентильных долей доходов может выявить особенности расслоения, которые другие методы, такие как медианный или средний доход, не могут полноценно отразить.

Для изучения распределения доходов с использованием метода процентилей используется следующая методология [3]:

1. Сбор данных: первым этапом является сбор данных о доходах. Это могут быть данные из национальных статистических служб, социологических опросов или крупных исследований, проводимых экономическими институтами.

2. Сортировка по доходам: далее доходы сортируются по возрастанию, после чего производится деление на процентильные группы. Для анализа децильного распределения доходов, например, общая совокупность разбивается на 10 частей, в каждой из которых содержится 10% населения.

3. Расчёт долей доходов: после сортировки рассчитываются суммарные доходы в каждом процентиле и их доля от общего дохода, это позволяет определить, какую часть совокупного дохода получают различные процентильные группы.

4. Графический анализ: на основании расчётов строится график процентильного распределения доходов, где на горизонтальной оси указаны процентильные группы, а на вертикальной доля доходов,

приходящаяся на каждую из них. Такой график позволяет наглядно увидеть, насколько концентрирован доход в верхних группах населения.

Примером применения процентильного подхода является децильный коэффициент Джини, показывающий соотношение доходов между различными децилями, а также анализ 90-го процентиля для оценки уровня доходов, получаемых наиболее обеспеченными группами населения. Высокий уровень концентрации доходов в верхних процентилях указывает на наличие значительной неравномерности в распределении, демонстрируя необходимость дополнительных мер со стороны государственной политики для снижения уровня социального расслоения [2].

Преимущества метода:

1. простота интерпретации: использование процентильных групп делает результаты анализа легко интерпретируемыми, что способствует их использованию в любой сфере.
2. выявление концентрации доходов: данный метод позволяет выявить группы, в которых концентрируется значительная доля доходов, что не всегда возможно с использованием других статистических методов.
3. сравнительный анализ: метод процентилей также полезен при сравнении доходных структур разных стран или регионов, он позволяет оценить, как распределяются доходы в различных обществах.

Метод процентилей, хотя и является информативным, имеет ряд ограничений:

1. невозможность выявления индивидуальных отклонений: процентильный подход рассматривает лишь общие группы, но не учитывает индивидуальные отличия;
2. сложность учёта внешних факторов: например, инфляция или различия в уровне жизни между регионами могут влиять на доходы, что затрудняет интерпретацию результатов;
3. ограниченная применимость в малых выборках: в малых выборках процентильный анализ может быть искажённым, так как выборки недостаточно информативны и полны для получения объективного представления о распределении доходов.

Анализ неравномерного распределения доходов с применением метода сортировки по процентиям является мощным инструментом для исследования экономического неравенства. Метод позволяет выявить ключевые особенности концентрации доходов и уровень социального расслоения в обществе, что делает его ценным для

экономистов, политологов и социологов. Однако для полноценного анализа рекомендуется использовать метод процентилей в совокупности с другими статистическими методами, позволяя получить более полное представление о состоянии социального неравенства.

Литература:

1. Atkinson A.B., Bourguignon F. Handbook of Income Distribution. Elsevier, 2014.
2. Cowell F. Measuring Inequality. Oxford University Press, 2011.
3. Deaton A. Analysis of Household Surveys. The World Bank, 1997.
4. Jenkins S.P. The Income Distribution in OECD Countries. Oxford Economic Papers, 2017.

РЕКОНСТРУКЦИЯ РАСТРОВЫХ ИЗОБРАЖЕНИЙ С УМЕНЬШЕННЫМ КОЛИЧЕСТВОМ ПИКСЕЛЕЙ

А.А. Упольников, Н.А. Кириллова
г. Новосибирск, sansan857@mail.ru

Научный руководитель: Кириллова Н.А., учитель математики
МАОУ «Лицей №126»

В работе рассматривается технология реконструкции изображения, позволяющая восстанавливать растровые изображения с уменьшенным количеством пикселей и с сохранением контрастности. На основе снимков фиксации нарушения правил дорожного движения продемонстрирована эффективность разработанного программного обеспечения.

The paper considers an image reconstruction technology that allows you to restore bitmap images with a reduced number of pixels and maintaining contrast. The effectiveness of the developed software is demonstrated on the basis of photographs of traffic violations.

Задача повышения качества и разрешения изображений стоит остро при получении изображений фото фиксирующими приборами и устройствами медицинской диагностики, при распознавании лиц и движущихся объектов, при распознавании фигур заданной формы и др. Так, например, при фото фиксации нарушения правил дорожного движения (ПДД), возможны ситуации с нечеткой идентификацией номеров, требующих дополнительной обработки (см. рис 1).



Рис. 1. – фото фиксация нарушения ПДД

В настоящее время для реконструкции изображений применяются следующие методы: линейные, нелинейные и методы суперразрешения. Основными линейными методами для увеличения масштаба изображения являются линейные интерполяторы, включающие в себя метод ближайшего соседа, билинейный метод и бикубический метод. Указанные методы широко используются из-за их простоты и не требовательности ресурсов. К нелинейным методам реконструкции можно отнести метод искаженного расстояния (WADI, от англ. warped distance), метод интерполяции (NEDI, от англ. new edge-directed interpolation). Методы суперразрешения требуют достаточного объема данных для обучения, на основе которых строится набор фильтров, в результате применения которых к изображениям, повышается их разрешение.

На сегодняшний день существуют графические редакторы, наиболее популярным из которых является Adobe Photoshop, использующий для ресаплинга изображений как стандартные линейные алгоритмы, так и улучшенные бикубический метод.

В работе разработано алгоритмическое и программное обеспечение для реконструкции растровых изображений с уменьшенным количеством пикселей. Разработанный алгоритм реконструкции изображений основан на методе билинейной интерполяции и позволяет на основе обучения сформировать фильтры, восстанавливающие снимки с повышением детализации и четкости снимка (см. рис 2,3).



Рис.2 – первоначальное изображение с камеры фото фиксации нарушения ПДД



Рис.3 - изображение после реконструкции

Литература:

1. Дьяконов В. MATLAB. Обработка сигналов и изображений. Специальный справочник / В. Дьяконов, И. Абраменкова. – СПб.: Питер, 2002. – 608 с.
2. Поршнев С.В. MATLAB 7 Основы работы и программирования / С.В. Поршнев. – М.: Бином-Пресс, 2010. – 320 с.
3. Крылов А.С. Компьютерное повышение разрешения изображений с использованием методов математической физики / А.С. Крылов, А.В. Насонов. – М.: МГУ им. М.В. Ломоносова, 2011. – 72 с.
4. Маркелов К.С. Модель повышения информативности цифровых изображений на базе метода суперразрешения [Электронный ресурс]: электронный научно-технический журнал. – /К.С. Маркелов. - Электрон. журн. – Москва: ФГБОУ ВПО «Московский государственный технический университет им. Н.Э. Баумана», 2013. – Режим доступа: <http://engsi.ru/doc/552065.html>.

ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ В АНАЛИЗЕ ЭКОНОМИЧЕСКИХ ДАННЫХ

А.О. Феденко, А.В. Фаддеенков

**Заполярный государственный университет им. Н.М. Федоровского.
г. Норильск, fao@bmit.ai**

В работе рассматривается применение методов машинного обучения для анализа и прогнозирования экономических данных. Изучены возможности алгоритмов регрессии, кластеризации и деревьев решений, используемых для предсказания рыночных трендов и анализа потребительского поведения. Подробно описана интеграция эконометрических моделей с методами машинного обучения для повышения точности анализа и прогнозирования в экономике.

The paper considers the application of machine learning methods for the analysis and forecasting of economic data. The possibilities of regression algorithms, clustering and decision trees used to predict market trends and analyze consumer behavior are studied. The integration of econometric models with machine learning methods to improve the accuracy of analysis and forecasting in economics is described in detail.

Современная экономика находится под влиянием множества факторов, изменяющихся в реальном времени: колебания цен, потребительские предпочтения, макроэкономические условия. Чтобы оставаться конкурентоспособными, компании и государственные организации стремятся к более точным экономическим прогнозам. Традиционные эконометрические методы, такие как регрессионный анализ и анализ временных рядов, полезны, но ограничены в гибкости и объеме данных, которые они могут обрабатывать. Машинное обучение, напротив, позволяет анализировать большие данные и выявлять сложные зависимости, что делает его привлекательным для экономического анализа.

В качестве алгоритмов машинного обучения в экономическом анализе обычно используют следующие методы:

1) Регрессионные методы широко применяются для анализа и прогнозирования экономических данных.

Наиболее распространенными типами регрессии являются:

1.1) Линейная регрессия: используется для анализа взаимосвязи между переменными. Примером может быть предсказание прибыли компании на основе изменения цен на сырьевые ресурсы.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon \quad (1)$$

где y — зависимая переменная (например, прибыль), x_1, x_2, \dots, x_n — независимые переменные (например, цены на сырьевые ресурсы), β_0 — свободный коэффициент, $\beta_1, \beta_2, \dots, \beta_n$ — коэффициенты, ϵ — ошибка.

1.2) Логистическая регрессия: помогает классифицировать данные, например, прогнозировать, дефолтный ли заемщик, исходя из его финансовой истории. Уравнение логистической регрессии:

$$P(y = 1 | X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \quad (2)$$

где $P(y = 1 | X)$ — вероятность того, что результат $y = 1$, x_1, x_2, \dots, x_n — независимые переменные, $\beta_0, \beta_1, \dots, \beta_n$ — коэффициенты.

1.3) Полиномиальная регрессия: подходит для анализа сложных зависимостей, например, для предсказания потребительских расходов в зависимости от дохода и других факторов. Уравнение полиномиальной регрессии:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \dots + \beta_n x^n + \epsilon \quad (3)$$

где n — степень полинома.

2) Кластеризация помогает сегментировать данные на группы, что особенно полезно в маркетинге и анализе потребительского поведения.

2.1) Алгоритм k-средних часто применяется для сегментации потребителей. Например, можно разделить клиентов на группы в зависимости от их покупательской способности, возраста и географии. Уравнение алгоритма k-средних:

$$J = \sum_{i=1}^m \sum_{k=1}^K 1(y^{(i)} = k) \|x^{(i)} - \mu_k\|^2 \quad (4)$$

где m — количество точек данных, K — количество кластеров, $x^{(i)}$ — точка данных, μ_k — центроид кластера k , $y^{(i)}$ — метка, указывающая на принадлежность точки к кластеру.

2.2) Иерархическая кластеризация помогает создать "дерево" групп, которое можно использовать для анализа иерархических зависимостей в больших данных. Уравнение иерархической кластеризации:

$$d(C_i, C_j) = \min_{a \in C_i, b \in C_j} \|a - b\| \quad (5)$$

где C_i, C_j — кластеры, a, b — элементы в кластерах.

3) Деревья решений и случайные леса - алгоритмы полезны для анализа рисков и создания прогнозов.

3.1) Деревья решений позволяют предсказывать вероятность определенного события на основе множества факторов. Используются для разделения данных на основе критериев (например, энтропии или коэффициента Джини). Например, они могут помочь оценить, как повлияет рост зарплат на потребительские расходы.

$$\text{Entropy}(S) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 - \dots \quad (6)$$

где p_1, p_2 — вероятности классов в выборке S .

3.2) Случайные леса объединяют несколько деревьев решений для повышения точности предсказаний. Часто используются для оценки риска или анализа кредитоспособности заемщиков.

Пример: банки применяют случайные леса для анализа вероятности дефолта клиентов, используя данные о доходах, кредитной истории и других факторах.

$$\tilde{y} = \frac{1}{T} \sum_{t=1}^T f_t(x) \quad (7)$$

где T — количество деревьев в лесу, $f_t(x)$ — предсказание для t -ого дерева.

Для иллюстрации применения линейной регрессии в анализе экономических данных рассмотрим задачу прогнозирования прибыли компании на основе изменения цен на сырьевые ресурсы. В данном примере используется уравнение линейной регрессии (1).

Компания стремится определить, как изменение цен на нефть и газ влияет на её прибыль. Исходные данные включают исторические значения цен на ресурсы и соответствующие показатели прибыли. Для построения модели линейной регрессии используется метод наименьших квадратов, который минимизирует сумму квадратов ошибок между фактическими и прогнозируемыми значениями прибыли.

На основании исторических данных была построена модель:

$$Y = 50 + 0.03X_1 + 0.02X_2$$

$b_0 = 50$, базовая прибыль компании в условиях, когда цены на ресурсы равны нулю.

$b_1 = 0.03$, каждый руб. роста цены нефти на 100 руб. увеличивает прибыль компании на 3 млн руб.

$b_2 = 0.02$, каждый руб. роста цены газ на 100 руб. увеличивает прибыль компании на 2 млн руб.

Предположим, что цена на нефть составляет 4600 руб./баррель, а цена на газ — 3100 руб./1000 куб. м. Тогда прогнозируемая прибыль будет рассчитана следующим образом:

$$Y = 50 + 0.03 * 4600 + 0.02 * 3100 = 50 + 138 + 62 = 250 \text{ млн. руб}$$

Прогноз показывает, как экономические параметры (цены на нефть и газ) влияют на финансовые результаты компании. Использование линейной регрессии в данном контексте позволяет: оценивать чувствительность прибыли компании к изменениям цен на сырье; проводить экономическое планирование на основе различных сценариев изменения рыночных условий; формировать стратегии ценообразования и управления затратами.

Машинное обучение (ML) связано с этим примером, так как алгоритм линейной регрессии является базовым методом ML. Его применение здесь заключается в:

1) Обучении модели: Алгоритм "учится" находить зависимости между входными (цены на нефть и газ) и выходными (прибыль) данными. Это обучение происходит на основе исторических данных.

2) Прогнозировании: после обучения модель может предсказывать новые значения прибыли при изменении цен на сырье. Это пример применения машинного обучения для анализа и прогнозирования.

3) Обобщении: Модель линейной регрессии не ограничивается только известными значениями данных. Она способна обобщать и предсказывать результаты для новых комбинаций факторов.

Линейная регрессия является базовым алгоритмом машинного обучения, который используется для анализа взаимосвязей между переменными и прогнозирования. В данном случае она демонстрирует свою эффективность для решения задачи экономического анализа, где требуется быстрое и точное предсказание на основе больших объемов данных. Кроме того, этот метод может быть интегрирован с более сложными моделями машинного обучения, такими как ансамбли или нейронные сети, для улучшения точности прогнозов в более сложных экономических сценариях.

Линейная регрессия является простой моделью, она формирует основу более сложных алгоритмов в ML, таких как полиномиальная регрессия, деревья решений, случайные леса и нейронные сети. Это объясняет, почему линейная регрессия рассматривается как часть машинного обучения, особенно в экономическом анализе, где требуется предсказание трендов и зависимостей.

Литература:

1. *Хасты Т., Тибширани Р., Фридман Дж.* Элементы статистического обучения: Выводы, прогнозирование и классификация. – М.: МЦНМО, 2015.

2. Goodfellow I., Bengio Y., Courville A. Deep Learning. – Cambridge: MIT Press, 2016.
3. Фишман Дж. Машинное обучение для бизнеса: Методы и примеры. – М.: Бомбора, 2018.
4. Tsay R. S. Analysis of Financial Time Series. – New York: Wiley, 2010.
5. Luca M., Bazerman M. H. The Power of Experiments: Decision Making in a Data-Driven World. – Cambridge: MIT Press, 2020.
6. Breiman L. Random Forests // Machine Learning. – 2001. – Vol. 45, No. 1. – URL: <https://link.springer.com/article/10.1023/A:1010933404324> (дата обращения: 23.11.2024).
7. Varian H. R. Big Data: New Tricks for Econometrics // Journal of Economic Perspectives. – 2014. – Vol. 28, No. 2. – URL: <https://pubs.aeaweb.org/doi/10.1257/jep.28.2.3> (дата обращения: 23.11.2024).

ОЦЕНКА РАССТОЯНИЙ МЕЖДУ АМИНОКИСЛОТНЫМИ ПОСЛЕДОВАТЕЛЬНОСТЯМИ НА ОСНОВЕ ИХ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ

К.О. Форкин, А.А. Зиянуров
Новосибирский государственный технический университет
Научный руководитель: Сивак М.А., к.т.н.

Задача оптимального выравнивания аминокислотных последовательностей является ключевой для анализа и предсказания их третичной структуры. В статье рассматривается подход, использующий машинное обучение, а именно модель ProtBERT, для создания векторных представлений аминокислот, что позволяет проводить более точные выравнивания.

The task of optimal alignment of amino acid sequences is crucial for analyzing and predicting their tertiary structure. This paper discusses an approach that uses machine learning, specifically the ProtBERT model, to create vector representations of amino acids, enabling more accurate alignments.

Задача поиска оптимального парного или множественного выравниваний аминокислотных последовательностей является одним из основных этапов построения или уточнения их третичной структуры. Оптимальным выравниванием аминокислотных последовательностей называют такое выравнивание, при котором достигается наилучшее соответствие между двумя или более последовательностями аминокислот, учитывая различные критерии, такие как количество

совпадающих аминокислот, количество вставок или удалений, и степень схожести между аминокислотами.

Оптимальные парные выравнивания часто строятся с использованием алгоритмов, например, Нидлмана-Вунша (для глобального выравнивания) или Смита-Ватермана (для локального выравнивания), которые используют матрицы аминокислотных замен и штрафы за разрывы.

Главный минус использования данного подхода заключается в том, что матрицы аминокислотных замен, такие как PAM и BLOSUM [1], построены на основе статистических данных о частоте замен аминокислот в определенных эволюционных контекстах. Иными словами, матрицы не учитывают порядок следования аминокислот в конкретной исследуемой последовательности. Например, аминокислота, которая находится в активном центре белка, может иметь более жесткие ограничения на замены, чем аминокислота в менее важном участке. Статистические матрицы не учитывают такие локальные структурные и функциональные особенности, что может приводить к неоптимальным выравниваниям [2].

Для получения более качественных выравниваний с учетом биологического контекста и эволюционных зависимостей активно применяются методы машинного обучения. Например, модель ProtBERT [3] позволяет преобразовать аминокислотные последовательности, которые являются последовательностью символов, в векторные представления, отражающие свойства аминокислот в контексте последовательности [2].

Для вычисления расстояния между векторными представлениями аминокислотных последовательностей используются разные метрики:

1. Косинусное расстояние (Cosine Distance)

$$d_{\cos}(A, B) = 1 - \frac{AB}{\|A\| \|B\|}, \quad (1)$$

где A, B – векторные представления аминокислот A и B.

2. Евклидово расстояние (Euclidean Distance)

$$d_{euclidean}(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2}, \quad (2)$$

где A_i , B_i – i -е координаты векторных представлений аминокислот A и B.

3. Манхэттенское расстояние (Manhattan Distance)

$$d_{manhattan}(A, B) = \sum_{i=1}^n |A_i - B_i|, \quad (3)$$

где A_i , B_i – i -е координаты векторных представлений аминокислот А и В.

Для получения выравниваний последовательностей алгоритмами Нидлмана-Вунша и Смита-Ватермана вместо матриц аминокислотных замен использовались матрица, составленная из расстояний между векторными представлениями каждой аминокислоты из последовательностей.

Полученные результаты сравнивались с выравниваниями, полученными стандартным способом с использованием матриц аминокислотных замен, и истинными выравниваниями, которые были взяты из открытых источников [4].

Исследования проводились на семействе белков cd19822 Vbox2_MID1_C-I. Для определения выравнивания, наиболее приближённого к истинному, использовалось расстояние Джакарда:

$$d_J(S_1, S_2) = \frac{|S_1 \cup S_2| - |S_1 \cap S_2|}{|S_1 \cup S_2|}, \quad (4)$$

где S_1, S_2 – парные выравнивания последовательностей.

В результате проведённого исследования выяснилось, что выравнивания, полученные с использованием векторных представлений аминокислотных последовательностей, полученных с помощью модели ProtBERT, оказались ближе к эталонным выравниваниям с точки зрения расстояния Джакарда по сравнению с выравниваниями, полученными стандартным способом с помощью матрицы аминокислотных замен BLOSUM. При этом между векторными представлениями аминокислот вычислялось косинусное расстояние. При использовании евклидова и манхэттенского расстояния выравнивание получалось хуже.

В заключении хотим отметить критически важную роль выравнивания аминокислотных последовательностей в предсказании и моделировании трехмерной структуры белков, процессе, известном как белковый фолдинг. Полученные выравнивания не просто предоставляют информацию о сходстве последовательностей, они служат основой для решения множества задач, связанных с определением пространственной структуры белка. Одна из самых важных задач – это построение трехмерной модели белка на основе его аминокислотной последовательности, когда информация о пространственном расположении атомов неполна или отсутствует вовсе. Это особенно актуально для белков, структура которых не определена экспериментально методами рентгеноструктурного анализа.

Литература:

1. *Огурцов А.Н.* Методы Биоинформационного анализа. - Харьков: НТУ «ХПИ», 2011. - 113 с.
2. *Sepehr Ashrafzadeh, G. Brian Golding, Silvana Ilie, Lucian Ilie.* Scoring alignments by embedding vector similarity [Electronic resource] // Oxford Academic – 2024, May – URL: <https://academic.oup.com/bib/article/25/3/bbae178/7660710> (accessed: 23.10.2024).
3. *Ahmed Elnaggar, Michael Heinzinger, Christian Dallago, Ghalia Rehawi, Yu Wang, Llion Jones.* ProtTrans: Toward Understanding the Language of Life Through Self-Supervised Learning // IEEE Xplore – 2021, July – URL: <https://ieeexplore.ieee.org/document/9477085/authors#authors> (дата обращения: 28.10.2024).
4. CDD Conserved Protein Domain Family [Electronic resource]. – URL: <https://www.ncbi.nlm.nih.gov/Structure/cdd/cd19823> (accessed: 05.10.2024).

СТАТИСТИЧЕСКИЙ АНАЛИЗ ВОЛОНТЕРСКОГО ДВИЖЕНИЯ В РОССИЙСКОЙ ФЕДЕРАЦИИ ЗА 2017-2023 ГГ.

М.В. Чернова

**Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, mashachernova3@gmail.com**

Научный руководитель: Шмарихина Е.С., к.э.н., доцент

В статье представлен статистический анализ волонтерской деятельности в России с 2017 по 2023 годы с акцентом на сезонность и структурные различия в возрастном распределении волонтеров. По данным Федеральной службы государственной статистики рассчитаны поквартальные индексы сезонности численности волонтеров. Анализ структурных сдвигов по возрасту и структурных различий по видам деятельности в 2023 году проведен с использованием индекса Рябцева.

The article presents a statistical analysis of volunteer activity in Russia from 2017 to 2023, with an emphasis on seasonality and structural differences in the age distribution of volunteers. According to the Federal State Statistics Service, quarterly indices of the seasonality of the number of volunteers are calculated. The analysis of structural shifts in age and structural differences by type of activity for 2023 carried out using the Ryabtsev index.

Согласно официальной статистической методологии формирования системы показателей трудовой деятельности, занятости и недоиспользования рабочей силы Росстата, трудовая деятельность –

любая деятельность, осуществляемая лицами любого пола и возраста в целях производства товаров и оказания услуг для использования другими лицами или для собственного использования. Отдельной формой трудовой является быстро распространяющаяся в последнее время волонтерская деятельность. Лица, работающие в качестве волонтеров, определяются как все лица обследуемого возраста, которые в течение короткого учетного периода (четыре недели или один календарный месяц) осуществляли любую неоплачиваемую, необязательную деятельность, связанную с производством товаров или оказанием услуг в интересах других лиц [1].

5 февраля 2018 года Президентом России был подписан федеральный закон "О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам добровольчества (волонтерства)" [2]. В этом документе уравниваются понятия волонтерской и добровольческой деятельности, определяются права и обязанности волонтеров, а также устанавливаются полномочия органов государственной власти в сфере добровольчества. Волонтерская деятельность в Российской Федерации в последние годы приобретает существенное значение, становясь неотъемлемой частью рынка труда и социальной сферы. В период с 2017 по 2023 годы наблюдается устойчивый рост вовлеченности населения в добровольческие инициативы. Поддержка добровольчества со стороны государства, общественных организаций и частного сектора побуждает все больше людей, независимо от возраста и профессии, заниматься общественно полезным трудом [3].

Выявление сезонности позволяет более точно прогнозировать численность волонтеров в течение года, что способствует эффективному распределению ресурсов и планированию мероприятий [4]. Для изучения сезонных колебаний численности волонтеров в период 2021-2023 годов (рисунок 1) был применен индексный метод, выявивший зависимость численности волонтеров от времени года.

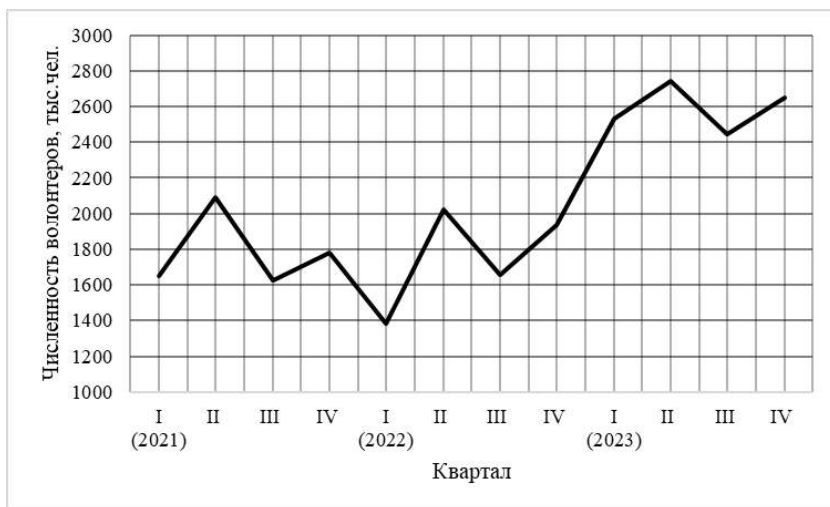


Рис.1 – Динамика численности волонтеров в России по кварталам за 2021-2023 гг.

Полученные результаты расчетов общего квартального индекса сезонности демонстрируют различия в активности волонтеров по кварталам:

$$I_{1\text{сез}} = 0,9759, I_{2\text{сез}} = 1,1454, I_{3\text{сез}} = 0,9127, I_{4\text{сез}} = 0,9717.$$

Эти данные показывают, что самый высокий уровень волонтерской активности приходится на второй квартал, на что, вероятно, влияет улучшение погодных условий и увеличение количества мероприятий, связанных с благоустройством и очисткой территорий населенных пунктов и других экологических проектов. Самый низкий уровень активности приходится на третий квартал, что, возможно, связано с сезоном отпусков. Сезонные колебания подчеркивают важность учета времени года для более успешной организации и привлечения волонтеров.

Анализ сезонных колебаний показывает значимость временного фактора при организации волонтерских мероприятий, также не менее важно понимать возрастную структуру волонтеров, поскольку возраст влияет на предпочтения и возможности участников в различных видах деятельности. Для оценки изменений в возрастной структуре волонтеров с 2017 по 2023 годы (таблица 1) использован индекс Рябцева, позволяющий измерить уровень разницы между возрастными структурами в два периода времени.

**Таблица 1 – Возрастная структура численности волонтеров
России в 2017 и 2023 гг.**

Возрастные группы, лет	Доля численности волонтеров данного возраста	
	2017	2023
15-19	0,083	0,113
20-29	0,161	0,158
30-39	0,188	0,193
40-49	0,167	0,174
50-59	0,195	0,157
60-69	0,147	0,139
70 и старше	0,060	0,066
Итого:	1,000	1,000

Индекс Рябцева составил 0,0633, что говорит о весьма низком уровне различий в возрастной структуре за последние шесть лет. Это показывает, что волонтерская деятельность в России с течением времени привлекает примерно одни и те же возрастные группы и их увеличение пропорционально, что указывает на стабильность возрастного состава волонтеров и сохранение интереса к волонтерству среди отдельных возрастных категорий.

Статистические данные о численности волонтеров представлены также по видам деятельности. Сравнивая попарно все виды деятельности в возрастном разрезе можно оценить структурные различия, также используя индекс Рябцева.

В 2023 году наименьшие различия в возрасте наблюдаются между волонтерами, занимающимися экологическими проектами (расчистка лесов, полей и рек, посадка деревьев), и волонтерами, занимающимися организацией спортивных, культурных и других общественных мероприятий, где индекс равен 0,0965. Среди волонтеров этих групп преобладает население в возрасте 20-29 лет, которое составляет 35,4% и 28,5%, соответственно. Также небольшие различия были обнаружены в возрастной структуре волонтеров, занимающихся сельскохозяйственными работами и сбором средств на благотворительность (индекс Рябцева - 0,0809). Больше всего данными видами волонтерской деятельности занимаются люди 30-39 лет, составляющие 28,6% и 24,4% от общего числа, соответственно. Можно сделать вывод, что данные направления популярны среди определенных возрастных групп, а именно, молодежи.

Наибольшие различия в возрастном составе были обнаружены между волонтерами, участвующими в экологических проектах и сельскохозяйственных работах (индекс - 0,4985). Несмотря на схожесть этих двух видов деятельности, молодёжь наиболее активно принимает участие в экологических проектах, а работа на фермах и в полях больше привлекает людей среднего возраста. Также высокий уровень возрастных различий между волонтерами, помогающими местным школам, и теми, кто предоставляет бесплатную медицинскую помощь или юридическую консультацию, проводит неоплачиваемые занятия, тренировки, консультирование (индекс - 0,4716). Здесь различия связаны со специфическими навыками, опытом и физическими требованиями, характерными для разных видов волонтерской деятельности.

В течение года численность людей, занимающихся волонтерской деятельностью, меняется, имеет место сезонность. Главными факторами, способствующими этим изменениям, являются погодные условия и время года. Численность волонтеров постоянно растет, причем равномерно по всем возрастным группам, что подчеркивает стабильность интереса к добровольчеству. По видам деятельности, как правило, возрастной состав волонтеров меняется, так как молодежь, люди среднего возраста и пожилые часто заинтересованы в разных проектах. Кроме того, возраст часто является барьером для занятия каким-либо видом волонтерства, к примеру, людям в старшем возрасте заниматься поиском пропавших людей может быть тяжело, а подростки не могут оказывать консультативную юридическую или медицинскую помощь из-за отсутствия соответствующего образования.

Литература:

1. Официальная статистическая методология формирования системы показателей трудовой деятельности, занятости и недоиспользования рабочей силы, рекомендованных 19-ой Международной конференцией статистиков труда от 31.12.2015 [Электронный ресурс] // Федеральная служба государственной статистики. – 2015. – URL: https://rosstat.gov.ru/labour_force (дата обращения: 01.11.2024).

2. Федеральный закон «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам добровольчества (волонтерства)» от 05.02.2018 № 15-ФЗ // Российская газета.

3. *Сергиевская И.А.* Волонтерство в контексте занятости // Экономика труда. – 2016. – №3. – С. 229–246.

4. Галочкин В.Т. Эконометрика: учебник и практикум для вузов. – М.: Юрайт, 2024. – 293 с.

НОВЫЙ СТАТИСТИЧЕСКИЙ МЕТОД ПРОВЕРКИ СТАЦИОНАРНОСТИ ВРЕМЕННЫХ РЯДОВ

Д.И. Чернышова

Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, beyourself20@mail.ru

Научный руководитель: Зотьев Д. Б., д. ф.-м. н., профессор

В настоящей статье представлен новый метод проверки слабой стационарности случайного процесса по его временному ряду. Научная новизна представленных в статье процедур состоит в объективно обоснованных оценках меры близости к нулю этих отличий, позволяющих принять гипотезу слабой стационарности или отвергнуть.

This article presents a new method for checking the weak stationarity of a random process over its time series. The scientific novelty of the procedures presented in the article consists in objectively justified estimates of the measure of proximity to zero of these differences, which make it possible to accept or reject the hypothesis of weak stationarity.

Случайным процессом (СП) называется зависящая от времени случайная величина $X(t)$. Важнейшей характеристикой СП является функция $m_X(t) = E(X(t))$. Линейная взаимосвязь случайных величин $X(t)$ и $X(t')$ характеризуется автоковариационной функцией (АКФ) [4]:

$$K_X(t, t') = E[(X(t) - m_X(t)) \cdot (X(t') - m_X(t'))] \quad \forall t \forall t' \quad (1)$$

Свойство слабой стационарности СП: значение $m_X(t)$ не зависит от t , а функция $K_X(t, t')$ (1) зависит только от $t' - t$, так что $K_X(t, t') = k_X(t' - t)$ для некоторой функции $k_X(\tau)$, где $k_X(\tau) = k_X(-\tau)$.

Для СП, представленного временным рядом x_i при $i = 0, 1, \dots, N$, условие слабой стационарности можно проверить следующим образом.

1с. Вычислим среднее значение уровней ряда \bar{x} . Рассматривая множество чисел $\{x_i: i = 0, 1, \dots, N\}$, как генеральную совокупность, оценим объем n выборки из этой совокупности, который с надежностью $P \geq 0,95$ обеспечивает погрешность ε среднего значения выборки \bar{x}_B по отношению к \bar{x} (предполагается, что \bar{x} не близко к нулю). Погрешность $|\bar{x} - \bar{x}_B|/|\bar{x}|$ можно оценить числом

$$\varepsilon = t_{(P+1)/2} \cdot \frac{\sigma}{|\bar{x}| \cdot \sqrt{n}} \quad (2)$$

где $t_{(P+1)/2}$ обозначает $(P+1)/2$ - квантиль распределения Стьюдента с $n-1$ степенями свободы [1]. Для корректного использования формулы (2) следует удалить из временного ряда x_i тренд. Таким образом, для случайной выборки объема n вероятность события $\{|\bar{x} - \bar{x}_b|/|\bar{x}| \leq \varepsilon\}$ равна P . При этом число n должно быть существенно меньшим, чем $N+1$, иначе почти стационарность величины \bar{x}_b будет тривиальным следствием ее близости к \bar{x} . Например, пусть $n \leq (N+1)/2$.

Для того, чтобы проверить стационарность среднего $m_X(t)$, мы вычислим $\bar{x}_b(m) = \sum_{j=m}^{m+n-1} x_j/n$ при каждом $m = 0, 1, \dots, N-n+1$ и найдем

$$\delta = \max_{0 \leq m \leq N-n+1} |\bar{x}_b(m) - \bar{x}|/|\bar{x}| \quad (3)$$

Если $\delta \leq \varepsilon$, то отклонения средних значений $\bar{x}_b(m)$ функции $x(t)$ при $m\Delta t \leq t \leq (m+n-1)\Delta t$ от ее среднего значения \bar{x} не превышают статистической погрешности оценки $\bar{x}_b \approx \bar{x}$, полученной из случайной выборки n уровней ряда. Таким образом, неравенства $\delta \leq \varepsilon$ подтверждают (хотя и не доказывают) гипотезу о том, что $m_X(t)$ не зависит от t . Если $\delta > \varepsilon$, но $\delta - \varepsilon \ll \delta$, то эту гипотезу также можно признать подтвержденной. Неравенство $\delta \gg \varepsilon$ следует считать свидетельством того, что $m_X(t)$ существенно зависит от t . Некоторые соотношения между δ и ε , например $\delta = 2\varepsilon$, попадают в «серую зону», из которой трудно извлечь определенное суждение.

2с. Для того, чтобы проверить стационарность АКФ $K_X(t_1, t_2)$, вычислим автоковариацию c_j^m для временного ряда $x_m, x_{m+1}, \dots, x_{m+n-1}$ при каждом $m = 0, 1, \dots, N-n+1$:

$$c_j^m = \frac{1}{n-j} \cdot \sum_{i=m}^{m+n-1-j} (x_i - \bar{x}_b(m)) \cdot (x_{i+j} - \bar{x}_b(m))$$

$$j = 0, 1, \dots, L \leq n/4 \quad (4)$$

Тогда $c_j^m \approx K_X(m\Delta t, (m+j)\Delta t)$. Вычислим автокоррелограмму $\rho_j^m = c_j^m/c_0^m$. Для любого фиксированного $j \in \{1, \dots, L\}$ число ρ_j^m - это выборочный коэффициент корреляции, который с вероятностью P попадает в доверительный интервал (5) [3, стр. 180]:

$$\left[\tanh \left(\operatorname{atanh}(\rho_j) + \frac{\rho_j}{2(n-1)} - \frac{u_{(P+1)/2}}{\sqrt{n-3}} \right); \right. \\ \left. \tanh \left(\operatorname{atanh}(\rho_j) + \frac{\rho_j}{2(n-1)} + \frac{u_{(P+1)/2}}{\sqrt{n-3}} \right) \right] \quad (5)$$

где $u_{(P+1)/2}$ – квантиль стандартного нормального распределения, отвечающая вероятности $(P + 1)/2$, и коэффициент корреляции $\rho_j = R_X(m\Delta t, (m + j)\Delta t)$. Каждый из промежутков (5) содержит в себе точку ρ_j^m с достаточно высокой вероятностью P . Гипотеза о стационарности автоковариационной функции $K_X(t_1, t_2)$ подтверждается, если при каждом $j = 1, \dots, L$ и каждом $m = 0, 1, \dots, N - n + 1$ выполнено следующее условие:

$$\begin{aligned} \tanh\left(\operatorname{atanh}(\rho_j) + \frac{\rho_j}{2(n-1)} - \frac{u_{(P+1)/2}}{\sqrt{n-3}}\right) &\leq \rho_j^m \leq \\ &\leq \tanh\left(\operatorname{atanh}(\rho_j) + \frac{\rho_j}{2(n-1)} + \frac{u_{(P+1)/2}}{\sqrt{n-3}}\right) \end{aligned} \quad (6)$$

Существенные нарушения неравенств (6) будут свидетельствовать о том, что АКФ нестационарна.

Для эффективного применения доверительного промежутка (5) коэффициент корреляции ρ_j должен быть близок по модулю к 1. Если это не так, то объем выборки n должен быть достаточно большим.

Применим метод проверки **1с** к временному ряду котировок акций компании «Инарктика» в течение 90 дней с 21.12.2022 по 20.03.2023 [2]. Здесь номер дня $j = 1, 2, \dots, 90$, моменты времени $t_{j-1} = j - 1$, а $\Delta t = 1$. Для того, чтобы сделать пример более выразительным, повысим вариативность данного ряда, отсчитывая котировки акций от минимального уровня 597 руб. Среднее значение уровней ряда $\bar{x} = 13,67$, стандартное отклонение $\sigma = 6,8$. Здесь $N + 1 = 90$, поэтому положим $n = 45$. Пусть $P = 0,95$. Оценим погрешность $|\bar{x} - \bar{x}_B|/|\bar{x}|$ согласно (1): $\varepsilon = t_{0,975} \cdot \sigma / (|\bar{x}| \cdot \sqrt{45}) = 0,15$, где $t_{0,975} \approx 2$. Согласно (2) $\delta = 0,128$. Поскольку $0,128 < \varepsilon$, то стационарность среднего значения $m_X(t)$ подтверждена по методу **1с**. Проверка стационарности АКФ по методу **2с** в данном случае не будет убедительной, т.к. объем выборки $n = 45$ слишком мал, и для большинства значений j модуль ρ_j далек от 1.

Рассмотрим модельный пример слабо стационарного процесса, который позволит нам проверить эффективность метода **2с**. Пусть Z – случайная величина, равномерно распределенная на отрезке $[0; 2\pi]$, $A \neq 0$, B и $\omega > 0$ – любые константы, случайный процесс $X_Z(t) = A \sin(\omega t + Z) + B$. Известно, что данный процесс слабо стационарен, $E(X_Z(t)) = B$ и $k_{X_Z}(\tau) = A^2/2 \cdot \cos(\omega\tau)$. Пусть $A = 1$, $B = 2$ и $\omega = 1$. Зафиксируем реализацию $x(t)$ данного процесса, которой отвечает значение $z = 0$. Тогда $x(t) = \sin(t) + 2$. Рассмотрим временной ряд x_i , являющийся дискретизацией функции $x(t)$, где $i = 0, 1, \dots, 499$ и $\Delta t =$

0,1. Пусть $N = 499$, $n = 100$ и $j = 5$. При каждом $m = 0, 1, \dots, 400$ условие (6) выполняется, что подтверждает стационарность по АКФ.

Теперь рассмотрим пример, в котором метод 2с дает отрицательный результат. Сделаем частоту ω зависящей от времени, так что $\omega = 1 + \sin(t)$. Вычисляя АКФ получим $K_{X_Z}(t_1, t_2) = A^2/2 \cdot \cos(\omega_2 t_2 - \omega_1 t_1)$. Как видно, функция $K_{X_Z}(t_1, t_2)$ не является стационарной. Рассмотрим реализацию $x(t) = \sin((1 + \sin(t)) \cdot t) + 2$. При $N = 499$, $n = 100$ и $j = 5$ неравенство (6) нарушается в 84-х случаях из 401-го. Это убедительно свидетельствует о том, что данный процесс не стационарен по АКФ.

В настоящем докладе представлен новый метод проверки слабой стационарности случайного процесса по его временному ряду. По существу они сводятся к проверке на стационарность среднего значения и автокорреляционной функции посредством их вычисления на временных рядах, являющихся отрезками исходного временного ряда. Научная новизна представленных в статье процедур состоит в объективно обоснованных оценках меры близости к нулю этих отличий, позволяющих принять гипотезу слабой стационарности или отвергнуть. Новые статистические критерии стационарности случайного процесса могут быть практически полезными. Однако, они нуждаются в дополнительной, эмпирической проверке.

Литература:

1. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Издательство Юрайт, 2024. – 479 с.
2. График Inarktika PAO (AQUA) [Электронный ресурс] / Investing.com. – URL: <https://ru.investing.com/equities/russian-sea-group-chart> (дата обращения: 23.10.2024).
3. Фёрстер Э., Рёнци Б. Методы корреляционного и регрессионного анализа. Руководство для экономистов. – М.: «Финансы и статистика», 1983. – 304 с.
4. Gimeno R., Manchado B., Mínguez R. Stationarity tests for financial time series // Physica A: Statistical Mechanics and its Applications, Elsevier. – 1999. – Vol. 269, N 1. – P. 72-78.

КОНЦЕПЦИЯ РАЗРАБОТКИ ИНТЕРАКТИВНОЙ ОБУЧАЮЩЕЙ ПРОГРАММЫ ПО ГЕОМЕТРИИ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ

К.И. Шайхисламова

**Муниципальное автономное образовательное учреждение «Лицей
№22 «Надежда Сибири»»,
г. Новосибирск, karina.shaykin@gmail.com**

В данной статье рассматривается концепция разработки интерактивной обучающей программы по геометрии с использованием технологий дополненной реальности (AR). Целью работы является создание образовательного инструмента, который поможет учащимся среднего звена лучше усваивать геометрические понятия через визуализацию и интерактивное взаимодействие. Обсуждаются теоретические основы, методы реализации и предполагаемая структура программы, а также ожидаемые результаты и преимущества использования AR в образовательном процессе.

This article discusses the concept of developing an interactive geometry learning programme using augmented reality (AR) technologies. The aim of the paper is to create an educational tool that will help secondary students to better learn geometric concepts through visualisation and interactive interaction. The theoretical background, implementation methods and intended structure of the programme are discussed, as well as the expected results and advantages of using AR in the educational process.

Геометрия является одной из ключевых дисциплин в школьном образовании, однако многие учащиеся сталкиваются с трудностями в восприятии абстрактных понятий. Традиционные методы обучения часто не обеспечивают достаточной наглядности, что затрудняет понимание пространственных отношений и свойств фигур. В связи с этим возникает необходимость в разработке инновационных подходов к обучению, которые могут повысить интерес учащихся и улучшить их понимание материала.

Технологии дополненной реальности (AR) предоставляют уникальные возможности для создания интерактивных обучающих приложений. AR позволяет визуализировать геометрические фигуры в реальном времени, что способствует лучшему пониманию и запоминанию материала.

Основная цель исследования — разработка концепции интерактивной обучающей программы по геометрии с использованием технологий дополненной реальности.

Для достижения этой цели были поставлены следующие задачи:

1. Изучить существующие методы преподавания геометрии и выявить их недостатки.
2. Определить возможности применения AR для улучшения учебного процесса.
3. Разработать структуру интерактивной программы, включающую элементы дополненной реальности.
4. Оценить потенциальные результаты использования данной программы в образовательном процессе.

В ходе исследования были проанализированы традиционные методы преподавания геометрии, включая использование учебников, чертежей и моделей. Выявлено, что многие учащиеся испытывают трудности при работе с абстрактными концепциями, что снижает их интерес к предмету.

Технологии AR позволяют создавать виртуальные объекты, которые могут быть интегрированы в реальную среду. Это дает возможность учащимся взаимодействовать с геометрическими фигурами, изменять их размеры, вращать и исследовать их свойства в трехмерном пространстве.

Программа будет включать следующие элементы:

- визуализация: отображение трехмерных геометрических фигур на экране устройства;
- интерактивность: возможность взаимодействия с фигурами (например, изменение размеров, вращение);
- задачи и викторины: включение интерактивных задач для проверки знаний и закрепления материала.
- обратная связь: предоставление рекомендаций и подсказок на основе действий учащихся.

Для реализации проекта по созданию интерактивной обучающей программы по геометрии с использованием технологий дополненной реальности можно использовать следующие инструменты и технологии:

1. Платформы для разработки:

- Unity: один из самых популярных игровых движков для разработки AR-приложений. Unity поддерживает множество платформ и имеет обширную библиотеку ресурсов и плагинов.
- Vuforia: платформа для дополненной реальности, работающая в связке с Unity, позволяющая отслеживать изображения и объекты в реальном времени.

2. Инструменты для 3D-моделирования:

- Blender: бесплатный инструмент для создания 3D-моделей и анимации;

- SketchUp: удобная программа для моделирования простых геометрических объектов.

3. Дополнительные платформы AR:

- ARKit (для iOS) и ARCore (для Android): платформы для разработки AR-приложений, предоставляющие инструменты для создания контента;

- MAXST SDK: набор инструментов для разработки AR-приложений с поддержкой маркерного отслеживания.

4. Мобильные устройства: для тестирования приложения потребуются смартфоны или планшеты с поддержкой AR.

Ожидается, что использование интерактивной обучающей программы по геометрии с элементами дополненной реальности позволит:

1. Повысить интерес учащихся к изучению геометрии.

2. Улучшить понимание пространственных понятий и свойств фигур.

3. Способствовать развитию критического мышления и навыков решения задач.

4. Создать условия для индивидуального подхода к обучению.

Создание интерактивной обучающей программы по геометрии с использованием технологий дополненной реальности представляет собой перспективное направление в образовательном процессе. Данная программа может значительно повысить эффективность обучения, сделать его более увлекательным и доступным для учащихся среднего звена. В будущем планируется проведение тестирования прототипа программы среди целевой аудитории для оценки ее эффективности и выявления возможных направлений для доработки.

Литература:

1. Kaufmann H., Schmalstieg D. Mathematics and geometry education with collaborative augmented reality // Computers & Graphics. - 2023. - Vol. 27. №3. - P. 339-345. - doi:10.1016/s0097-8493(03)00028-1

2. Маркерная технология дополненной реальности как средство развития геометрической грамотности: эффективность и когнитивная нагрузка / Гаврилова Т. А., Бажина П. С., Ходченко А. К., Замаева Я. А. // Перспективы науки и образования. - 2022. - № 6 (60). - С. 535-553. -doi: 10.32744/pse.2022.6.32

СОДЕРЖАНИЕ

НАУЧНОЕ НАПРАВЛЕНИЕ – ИНФОРМАТИКА, АВТОМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ И ИЗМЕРИТЕЛЬНАЯ ТЕХНИКА

Секция АВТОМАТИКА, ИЗМЕРЕНИЯ И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

Волобуев К.А., Катасонов Д.Н., Трубин В.Г.	Новосибирск	3
Цирульников К.А., Катасонов Д.Н.	Новосибирск	8
Ребус И.М.	Новосибирск	12
Каштанов А.А., Рева И.Л., Фурьев И.А., Каштанова Е.В.	Новосибирск	16
Парьгин Р.П.	Новосибирск	20
Абдукаримов Т.А., Ковальчук А.В.	Новосибирск	25
Куршин В.М., Чириков М.А.	Новосибирск	28
Хозяева В.Д., Чириков М.А.	Новосибирск	33
Патрушев Б.Л.	Новосибирск	37
Серебрянников В.А.	Новосибирск	38
Галашина А.В.	Новосибирск	41

Секция ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Перевалов Г.Г., Блинников Н.А.	Новосибирск	45
Галета Н.Е.	Новосибирск	49
Панарин С.В.	Новосибирск	53
Чуркина Л.И.	Новосибирск	56
Денисов М.О., Яковина И.Н.	Новосибирск	61
Хабаров С.М., Яковина И.Н.	Новосибирск	65
Масленников И.Г.	Новосибирск	70
Мельников С.К., Шахрай В.А.	Новосибирск	74
Мельников С.К., Шахрай В.А.	Новосибирск	77
Ларионова А.А.	Новосибирск	81
Белов А.Ю., Яковина И.Н.	Новосибирск	84
Рясной М.В., Яковина И.Н.	Новосибирск	88
Кочнев Д.К., Осипенко И.В., Яковина И.Н.	Новосибирск	92
Жакупов Р.А.	Томск	95
Каталицкий Г.А.	Новосибирск	98

Еремеева В.В., Бурлаков Н.Г.	Новосибирск	103
Гребенкин А.С.	Новосибирск	107
Скворцов М.В.	Новосибирск	111
Воробьева Ю.А., Уварова Е.А., Потапова И.А.	Санкт-Петербург	115
Кабелин М.Д., Моисеев Д.А.	Новосибирск	119
Галёб С., Климова Д.Н.	Ростов-на-Дону	124

**Секция ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЦИФРОВОЙ
ЭКОНОМИКЕ, УПРАВЛЕНИИ И В МЕДИЦИНЕ**

Апасова Д.С.	Новосибирск	128
Давыдов М.С.	Новосибирск	132
Захаров Д.М.	Ростов-на-Дону	137
Константинов Д.С.	Новосибирск	144
Лесик В.Е., Никифоров Д.В.	Москва	147
Озольс П.Ю.	Новосибирск	152
Яковлев Е.В.	Ростов-на-Дону	156

**НАУЧНОЕ НАПРАВЛЕНИЕ – ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ
И ОБРАБОТКИ ДАННЫХ**

**Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, АНАЛИЗ
И ОБРАБОТКА ДАННЫХ**

Шнайдер К.А.	Новосибирск	161
Ковригин Д.Д., Потапова И.А., Яковлева Т.Ю.	Санкт-Петербург	165
Осинцева Е.А.	Новосибирск	170
Комарова А.В.	Новосибирск	177
Пляскин Д.А.	Томск	182
Ражев К.С., Токарева А.В.	Томск	186
Потапова В.Р., Дьяченко Н.В., Михтеева Е.Ю.	Санкт-Петербург	191
Евланова С.А.	Новосибирск	198
Мадияров К.Г.	Новосибирск	202
Ефиркин М.И., Дыминский И.И.	Новосибирск	206
Левковский К.А.	Новосибирск	210
Пятанин А.А.	Новосибирск	213
Бурлаков В.С., Ишутин М.А.	Новосибирск	217

**Секция ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ
ПРОЦЕССОВ В ТЕХНОЛОГИЯХ И ПРИРОДНЫХ ЯВЛЕНИЯХ**

Булатова А.З., Биккинина Н.Б., Солнышкина О.А.	Уфа	222
Васильев Е.В.	Новосибирск	225
Виноградов К.М.	Новосибирск	230
Дербышева Т.Р.	Новосибирск	234
Иванов М.В., Персова М.Г.	Новосибирск	239
Исакин Д.А.	Новосибирск	243
Кравец Е.А., Персова М.Г.	Новосибирск	247
Лысенко Н.К.	Томск	252
Чмиленко Ф.В., Растворова И.И., Парменов В.Е.	Санкт-Петербург	256
Романов М.Д.	Новосибирск	261
Слобода В.С.	Новосибирск	265

**Секция ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА,
МАШИННОГО ОБУЧЕНИЯ И ПРЕДСКАЗАТЕЛЬНОГО
МОДЕЛИРОВАНИЯ**

Бессонов Н.А.	Ростов-на-Дону	270
Корнева Е.С.	Новосибирск	275
Кривецкий А.А., Карманов В.С.	Новосибирск	278
Мадияров К.Г.	Новосибирск	282
Мионов Л.А.	Новосибирск	286
Попова Д.А., Ткачук Ф.А.	Новосибирск	290
Самойленко Р.В.	Новосибирск	294
Собешиков Я.Ю.	Новосибирск	297
Филипенко Ю.Д.	Новосибирск	303
Хайдуков А.П.	Новосибирск	307
Чеглов Е.Р.	Новосибирск	311
Шайхисламов И.М.	Новосибирск	315

Секция ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ,
СТАТИСТИКА И ЭКОНОМЕТРИКА

Белозубов Г.В.	Новосибирск	320
Бортников Н.Е.	Новосибирск	323
Буксиков А.А., Фаддеенков А.В.	Норильск	328
Гафарова Л.О.	Новосибирск	332
Громов В.А., Фаддеенков А.В.	Норильск	336
Дворецкая В.К.	Новосибирск	339
Ерощенко А.А.	Новосибирск	344
Железнякова А.Д.	Новосибирск	349
Ицкович А.А.	Новосибирск	353
Кривецкий А.А., Карманов В.С.	Новосибирск	356
Марчев Е.Е., Плескач Д.М., Фаддеенков А.В.	Норильск	361
Мусаткин И.Д.	Новосибирск	363
Обидин Н.В.	Новосибирск	367
Рубан В.М.	Новосибирск	371
Рыбачёк Д.А.	Новосибирск	375
Салимзянов А.Н., Кобылянский В.Г.	Новосибирск	379
Семенов Т.В., Фаддеенков А.В.	Норильск	383
Упольников А.А., Кириллова Н.А.	Новосибирск	386
Феденко А.О., Фаддеенков А.В.	Норильск	389
Форкин К.О., Зиянуров А.А.	Новосибирск	393
Чернова М.В.	Новосибирск	396
Чернышова Д.И.	Новосибирск	401
Шайхисламова К.И.	Новосибирск	405

НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ
Сборник научных трудов в 8 частях

г. Новосибирск, 02 декабря - 06 декабря 2024 г.

Часть 1

Под редакцией Захаровой Е. В.

Подписано в печать 04.02.2025. Формат 60×84/16. Бумага офсетная.

Уч.-изд. л. 23,94. Печ. л. 25,75. Тираж 100 экз. Заказ №Р-04082

Отпечатано в типографии

Новосибирского государственного технического университета

630073, г. Новосибирск, пр. К. Маркса, 20