Министерство образования и науки РФ
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)

XII НАУЧНО- ПРАКТИЧЕСКАЯ КОНФЕРЕНЦИЯ

С МЕЖДУНАРОДНЫМ УЧАСТИЕМ

«НАУКА НАСТОЯЩЕГО И БУДУЩЕГО»

ДЛЯ СТУДЕНТОВ, АСПИРАНТОВ И МОЛОДЫХ УЧЕНЫХ

Сборник материалов конференции

15 – 17 мая 2025

Том I

Санкт-Петербург
2025

УДК 001.2

XIII НАУЧНО-ПРАКТИЧЕСКАЯ КОНФЕРЕНЦИЯ С МЕЖДУНАРОДНЫМ УЧАСТИЕМ «НАУКА НАСТОЯЩЕГО И БУДУЩЕГО» ДЛЯ СТУДЕНТОВ, АСПИРАНТОВ И МОЛОДЫХ УЧЕНЫХ. Том 1. Сборник материалов конференции. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2025. 258 с.

Организаторы:
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Тематика конференции включает следующие направления

- *Информационные радиотехнические системы и устройства*
- *Информатика и управление в технических системах и ВТ*
- *Программная инженерия и автономные интеллектуальные системы*
- *Искусственный интеллект в прикладных областях*
- *Алгоритмическая математика*
- *Управление и обработка информации*
- *Приборостроение*
- *Лингвистика*
- *Электроника, нанотехнологии, наноматериалы*
- *Системный анализ и информационная безопасность*
- *Электропривод, автоматика и электротехнологии*
- *Биотехнические системы и технологии*
- *Техносферная безопасность*
- *Реклама и связи с общественностью*
- *Современные тренды управления качеством и цифровая экономика,*
- *Инновационное проектирование: от реальных объектов к цифровым двойникам.*

Сборник материалов содержит доклады, представленные на XIII Научно-практической конференции с международным участием «Наука настоящего и будущего» для студентов, аспирантов и молодых ученых, состоявшейся 15 – 17 мая 2025 года в Санкт-Петербурге. Основной задачей конференции является развитие творческой активности студентов, привлечение их к решению актуальных задач в области науки и техники. Все доклады проходят рецензирование.

ОГЛАВЛЕНИЕ

СЕКЦИЯ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ПРИКЛАДНЫХ ОБЛАСТЯХ..... 8

Базарбаев К.А., Маннанов У.В.
ФОРМИРОВАНИЕ МНОГОМЕРНОГО УПРАВЛЕНИЯ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ..... 10

Абдулаев А.Р.
АНАЛИЗ И ПРОВЕРКА ТЕКСТА НА НАЛИЧИЕ ИСКУССТВЕННО СГЕНЕРИРОВАННЫХ ЧАСТЕЙ С ПОМОЩЬЮ LLM..... 14

Будило Е.А.
КЛАСТЕРНЫЙ АНАЛИЗ ПРОФЕССОРСКО-ПРЕПОДАВАТЕЛЬСКОГО СОСТАВА С ИСПОЛЬЗОВАНИЕМ NLP..... 18

Волкова А.А., Дружинская Е.В.
ОБЗОР МОДЕЛЕЙ АВТОМАТИЧЕСКОГО РАСПОЗНАНИЯ АКЦЕНТИРОВАННОЙ РЕЧИ 22

Горбачёв М.Е.
РАЗРАБОТКА КОМПАКТНЫХ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ДЛЯ ЛОКАЛЬНОЙ ОБРАБОТКИ ДАННЫХ В УСТРОЙСТВАХ ИНТЕРНЕТА ВЕЩЕЙ (IOT)..... 25

Жук Ф.П.
СЕГМЕНТАЦИЯ ГИСТОЛОГИЧЕСКИХ ИЗОБРАЖЕНИЙ СОСУДОВ ЛЕГКИХ НА ОСНОВЕ ГЛУБОКОГО ОБУЧЕНИЯ 27

Здех К.В, Викснин И.И.
ПОСТРОЕНИЕ ТРАЕКТОРИИ ПЕРЕМЕЩЕНИЯ ОБЪЕКТА НА ОСНОВЕ АНАЛИЗА ВИДЕОПОТОКА . 31

Иванова М.К.
КЛАССИФИКАЦИЯ МУЗЫКИ ПО ЭМОЦИЯМ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ..... 35

Иманбаев А.
ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ПРОГНОЗИРОВАНИЯ КАЧЕСТВА ВОЗДУХА НА ОСНОВЕ ГИБРИДНЫХ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ 39

Корноухова Т.В., Копылова Д.А.
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ОБУЧЕНИИ: ЛУЧШИЕ КЕЙСЫ ВНЕДРЕНИЯ В ШКОЛАХ И ВУЗАХ 41

Трофимов Ю.В, Кузнецов Е.М, Муравьев И.П.
МЯГКИЕ ИЗМЕРЕНИЯ НЕЙРОНЕЧЁТКИХ СОСТАВЛЯЮЩИХ XAI 2.0: МУЛЬТИУРОВНЕВАЯ ГИБРИДНАЯ АРХИТЕКТУРА (CNN + ANFIS) 44

Левушкина С.И, Панфилова Е.Д.
ОЦЕНКА ПРОИЗВОДИТЕЛЬНОСТИ И ЭФФЕКТИВНОСТИ ДИСТИЛИРУЕМЫХ, КВАНТУЕМЫХ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ С ОТКРЫТЫМ ИСХОДНЫМ КОДОМ ДЛЯ РАЗРАБОТКИ АГЕНТОВ ИИ НА ВСТРАИВАЕМЫХ УСТРОЙСТВАХ..... 47

Литягин С.М.
АНАЛИЗ СРЕД ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ МНОГОАГЕНТНЫХ СИСТЕМ РЕАЛЬНОГО ВРЕМЕНИ 52

Михайлова М.А.
ТЕХНОЛОГИЯ ОЦЕНКИ КАЧЕСТВА ЭЛЕКТРОТЕХНИЧЕСКИХ ИЗДЕЛИЙ МЕТОДАМИ МАШИННОГО ЗРЕНИЯ 56

Трофимов Ю.В, Кузнецов Е.М, Муравьев И.П.
МНОГОУРОВНЕВЫЙ ОБЪЯСНИТЕЛЬНЫЙ ИИ ДЛЯ ДИАГНОСТИКИ ПНЕВМОНИИ В РЕНТГЕН-СНИМКАХ: ОТ GRAD-CAM И LIME ДО ГЕНЕРАТИВНОГО АВТООПИСАНИЯ С MLOPS-ИНТЕГРАЦИЕЙ 60

Новицкий М.Д., Грибов А.В.
АНАЛИЗ МЕТОДОВ ПОИСКА ПУТИ ПОДВИЖНОГО АГЕНТА В ДИНАМИЧЕСКОЙ СРЕДЕ С ИСПОЛЬЗОВАНИЕМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ 63

Прахов Д.А.
ПРЕИМУЩЕСТВА И НЕДОСТАТКИ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ МЕДИЦИНСКОЙ ДИАГНОСТИКИ 67

Троицкая А.М., Фабричева Е.Д., Чепасов Д.В.
ПРИМЕНЕНИЕ 4D ГАУССОВСКОГО СПЛАТТИНГА ДЛЯ МОДЕЛИРОВАНИЯ ДИНАМИЧЕСКОЙ ТРЁХМЕРНОЙ СЦЕНЫ 71

Угольникова Е.А.
АВТОМАТИЗИРОВАННЫЙ РАСЧЕТ ПОКАЗАТЕЛЕЙ САГИТТАЛЬНОГО БАЛАНСА ПО ТРЕХМЕРНОЙ МОДЕЛИ ПОЗВОНОЧНИКА, ПОСТРОЕННОЙ ИЗ РЕНТГЕНОГРАММ С ПОМОЩЬЮ АЛГОРИТМОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ 75

Харитонов М.А.
АРХИТЕКТУРА СОВРЕМЕННЫХ СИСТЕМ НА ОСНОВЕ LLM 78

Чернов Е.А., Мелдо А.А., Агаева А.В., Богданов М.А., Вальков М.Ю
СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ВЫЖИВАЕМОСТИ ПРИ РАКЕ МОЛОЧНОЙ ЖЕЛЕЗЫ 82

Чепасов Д.В., Щедрин А.А.
РЕШЕНИЕ ЗАДАЧИ АВТОМАТИЧЕСКОГО ГЛОССИРОВАНИЯ ПРИ МИНИМАЛЬНОМ ОБЪЁМЕ ДАННЫХ: ПРИМЕР ИЖОРСКОГО ЯЗЫКА 85

СЕКЦИЯ АЛГОРИТМИЧЕСКАЯ МАТЕМАТИКА 89

Арутюнян С. Н.
МОДЕЛИРОВАНИЕ КОРОНАЛЬНЫХ ВЫБРОСОВ МАССЫ 89

Гаврилова О.П.
МОДЕЛИРОВАНИЕ КОРОНАЛЬНЫХ ВЫБРОСОВ МАССЫ ТРОПИЧЕСКАЯ КРИПТОГРАФИЯ. РЕАЛИЗАЦИЯ АЛГОРИТМА ЦЭП В ТЕРМИНАХ ТРОПИЧЕСКОЙ АЛГЕБРЫ 78

Кодуков А.В.
ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ ПОДХОДОВ В КЛАССИЧЕСКИХ АЛГОРИТМАХ СТЕРЕОЗРЕНИЯ ПРИ НЕДОСТАТКЕ ИНФОРМАЦИИ 98

Круглова В. Д.
ПОИСК ЦЕНТРАЛЬНЫХ МЕР НА ГРАФАХ ДВУСТРОЧЕЧНЫХ И ТРЕХСТРОЧЕЧНЫХ ДИАГРАММ ЮНГА 102

Найдёнова Е. В.
РАЗРАБОТКА МЕТОДОВ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ СТАТИСТИЧЕСКОЙ ОЦЕНКИ СЕЛЕКТИВНОСТИ ЗАПРОСОВ В РЕЛЯЦИОННЫХ СУБД С ИСПОЛЬЗОВАНИЕМ КОРРЕЛЯЦИОННОГО АНАЛИЗА 106

Субботин М.О.
ВЫЧИСЛЕНИЕ ОРБИТЫ КА MESSENGER ПО РАДИОТЕХНИЧЕСКИМ ИЗМЕРЕНИЯМ ДОПЛЕРОВСКОГО СДВИГА ЧАСТОТЫ В X-ДИАПАЗОНЕ 109

СЕКЦИЯ ИНФОРМАТИКА И УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ И ВТ 113

Вздорова М.А.
РАЗРАБОТКА АЛГОРИТМА ЭФФЕКТИВНОЙ ЦИФРОВИЗАЦИИ ПРОФИЛАКТИЧЕСКИХ ОСМОТРОВ В ВЕДОМСТВЕННЫХ ПОЛИКЛИНИКАХ МВД И МЧС 113

Генералов Г. Ю.
ТЕХНИЧЕСКИЕ АСПЕКТЫ СТЕНДА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ПОЖАРНОЙ СИГНАЛИЗАЦИИ, КОНТРОЛЯ ЗАГАЗОВАННОСТИ И ПОЖАРОТУШЕНИЯ..... 116

Здех С.А.
РЕАЛИЗАЦИЯ БИТ-ПОТОКОВОГО ПЕРСЕПТРОНА НА ПЛИС..... 119

Ковальчук И.А
КОНТРОЛЛЕР МОДУЛЯ ФОТОБИОМОДУЛЯЦИИ 123

Котов И. И., Куприянова А. М.
АВТОМАТИЗАЦИЯ ETL-ПРОЦЕССОВ С ИСПОЛЬЗОВАНИЕМ APACHE AIRFLOW 127

Куприянова А. М., Котов И. И., Куприянов Н. М.
АВТОМАТИЗАЦИЯ А/Б-ТЕСТИРОВАНИЯ 131

Мавлявиев Т. И., Моисеев К. М
РАЗРАБОТКА СИСТЕМЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ ТЕМПЕРАТУРОЙ ПОДЛОЖКОДЕРЖАТЕЛЯ НА УСТАНОВКЕ ВУП-11М 135

Макаров М.В.
ОРГАНИЗАЦИЯ ВЗАИМОДЕЙСТВИЯ С ДАТЧИКАМИ МОНИТОРИНГА ДОБЫВАЕМОЙ СМЕСИ НЕФТЕДОБЫВАЮЩИХ СКВАЖИН 139

Панфилова Е.Д.
УСТРОЙСТВО ЦИФРОВОГО МОНИТОРИНГА ТЕМПЕРАТУРЫ 144

Рогозинская А.Л.
ПРИМЕНЕНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ГЕМАТОЛОГИИ 148

Хасанов А.
РАЗРАБОТКА АЛГОРИТМА ЛОКАЛИЗАЦИИ ИСТОЧНИКА ЗАГРЯЗНЕНИЯ РАСПРЕДЕЛЕННЫМ РОБОТОТЕХНИЧЕСКИМ КОМПЛЕКСОМ НА ОСНОВЕ МУЛЬТИАГЕНТНОГО ПОДХОДА 150

Хорошков В.С., Казакова Т.Е., Колесникова А.В.
РОБОТ-СБОРЩИК КУБИКА РУБИКА 154

Черникова П.В.
ТЕХНОЛОГИЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ВИЗУАЛЬНОГО КОНТРОЛЯ КАЧЕСТВА ЭЛЕКТРОТЕХНИЧЕСКИХ ИЗДЕЛИЙ 159

Шевченко А.А.
ПОСТАНОВКА ПРАКТИКИ ПО РАЗРАБОТКЕ ВЕБ-СИСТЕМЫ ДЛЯ МОНИТОРИНГА ОБЪЕКТОВ И УПРАВЛЕНИЯ ИМИ 162

СЕКЦИЯ СИСТЕМНЫЙ АНАЛИЗ И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ 167

Буравлев А.С., Суслова К.А., Прохоров А.С.
ИДЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЕЙ ПО ГЕОМЕТРИИ РУКИ: МЕТОДЫ И ПЕРСПЕКТИВЫ 171

Иконников С.Н., Шкляр Е.В.
МЕТОДЫ ОПРЕДЕЛЕНИЯ ВИТАЛЬНОСТИ В БИОМЕТРИЧЕСКИХ СИСТЕМАХ 175

Катрущенко О.Д.
ПРОГРАММНЫЙ МОДУЛЬ ОПРЕДЕЛЕНИЯ ОШИБОК ПРИ РАЗРАБОТКЕ ПОЛЬЗОВАТЕЛЬСКИХ ИНТЕРФЕЙСОВ 179

Куприянов Н. М.
ПОДХОД К МОДЕЛИРОВАНИЮ ОТКАЗОУСТОЙЧИВЫХ ИНФОРМАЦИОННЫХ СИСТЕМ 183

Мерзляков И.А., Синев В.Е.
РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ УЧЁТА ПОТРЕБЛЕНИЯ КАЛОРИЙ С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА 187

Моисеев К.О., Трофимова А.О.
АНАЛИЗ ЭЛЕМЕНТОВ НАВИГАЦИИ СТАНЦИИ МЕТРО "МОСКОВСКАЯ" ПЕТЕРБУРГСКОГО МЕТРОПОЛИТЕНА 187

Налимов К.Г.
МЕТОДИКИ МОДЕЛИРОВАНИЯ УГРОЗ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В СИСТЕМАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА 195

Семейкин С.А., Кулагин М.В.
МЕТОДЫ ПОВЫШЕНИЯ КАЧЕСТВА ОПИСАНИЙ УЯЗВИМОСТЕЙ ПО НА ОСНОВЕ ДАННЫХ ИЗ ОТКРЫТЫХ ИСТОЧНИКОВ 198

Солонухина А.Л., Шкляр Е.В.
СТЕГАНОГРАФИЯ В НЕСЖАТОМ ВИДЕОПОТОКЕ С ИСПОЛЬЗОВАНИЕМ ДИСКРЕТНОГО КОСИНУСНОГО ПРЕОБРАЗОВАНИЯ 201

СЕКЦИЯ ИНФОРМАЦИОННЫЕ РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ И УСТРОЙСТВА 206

Барвиненко Д.С., Лысенко Н.В., Федотов Е.В.
КРЕМНИЕВЫЕ ФОТОПРИЕМНИКИ В ЗАДАЧАХ ТЕМПЕРАТУРНОГО КОНТРОЛЯ В МЕТАЛЛУРГИИ 206

Беспалов В.А., Бузуверов Г.В., Андреева О.М.
ОРГАНИЗАЦИЯ ОБМЕНА ДАННЫМИ В УСТРОЙСТВЕ РЕГИСТРАЦИИ СИГНАЛОВ АКУСТИЧЕСКОГО ДАТЧИКА 210

Зинченко А.А.
ЧАСТОТНАЯ СЕЛЕКЦИЯ ДВИЖУЩИХСЯ ЦЕЛЕЙ В РАДИОЛОКАЦИОННЫХ СИСТЕМАХ СО СЛОЖНЫМ СИГНАЛОМ 214

Гоношилкин Д. Ю., Баранов П. С.
МЕТОД КОМПЕНСАЦИИ ХРОМАТИЧЕСКИХ АБЕРРАЦИЙ ПРИ ФОРМИРОВАНИИ ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ, ПОЛУЧЕННЫХ С ПОМОЩЬЮ ПЕРЕСТРАИВАЕМОГО ИСТОЧНИКА СВЕТА 217

Гайворонский Д. В., Дианов А. В.
ИССЛЕДОВАНИЕ СТРУКТУРЫ СИГНАЛОВ С ПОМОЩЬЮ ПРЕОБРАЗОВАНИЙ ВЕЙВЛЕТОВ, ВИГНЕРА-ВИЛЯ И РАДОНА 221

Камели А. Н., Векшин Ю. В.
ИССЛЕДОВАНИЕ СТАБИЛЬНОСТИ ПРИЕМНИКА РАДИОТЕЛЕСКОПА В РЕЖИМАХ ПОЛНОЙ МОЩНОСТИ И МОДУЛЯЦИОННОМ 224

Ковалев Д.М., Сергиенко А.Б.
ПОМЕХОУСТОЙЧИВОСТЬ ПРИЕМА OTFS-СИГНАЛОВ В УСЛОВИЯХ БЫСТРЫХ ЗАМИРАНИЙ 228

Михайлов Д.В.
СРАВНИТЕЛЬНЫЙ АНАЛИЗ АЛГОРИТМОВ РАСПОЗНАВАНИЯ ТИПА ВНУТРИИМПУЛЬСНОЙ МОДУЛЯЦИИ В СРЕДСТВАХ РАДИОМОНИТОРИНГА 233

Хмыров И.К.
ОБЗОР СПОСОБОВ ПОСТРОЕНИЯ РЛС С СДЦ 237

Шатковская М.С.
ГИПЕРСПЕКТРАЛЬНАЯ СИСТЕМА ДЛЯ ДИАГНОСТИКИ ПОВРЕЖДЕНИЙ ТКАНЕЙ 240

Маноенко Г.С.
ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА МОДУЛЬНОГО ДРОНА С ОТКРЫТЫМ ПО 245

СЕКЦИЯ ЛИНГВИСТИКА 249

Александров К.С., Васильева О.А., Балаева К.А.
ОСОБЕННОСТИ БРИТАНСКОГО ЮМОРА 249

Кущенко А.П., Малышева В.Н.
МОДЕЛИ ОБРАЗОВАНИЯ СЛЕНГИЗМОВ В АМЕРИКАНСКОМ КИНОТЕКСТЕ: ЛИНГВИСТИЧЕСКИЙ И СОЦИОКУЛЬТУРНЫЙ АСПЕКТ 253

ПРИВЕТСТВЕННОЕ ВЫСТУПЛЕНИЕ

профессора Маннанова У.В. на
открытии XIII Всероссийской научно-практической конференции
«Наука настоящего и будущего»

Уважаемые коллеги,
уважаемые участники и гости конференции!

Мне доставляет огромное удовольствие приветствовать открытие Тринадцатой Всероссийской научно-практической конференции «Наука настоящего и будущего»!

Уже стало традицией ежегодно на эту конференцию с международным участием приглашать школьников, студентов, аспирантов, молодых ученых из многих стран, в том числе и из Узбекистана!

Мы гордимся этим, мы рады предоставленной возможности принять участие в такой конференции знаменитейшего высшего технического университета, мы рады, что в истории становления и развития Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» имени В.И.Ульянова есть пусть малая, но уверен, весомая толика и Ташкентского государственного технического университета имени Ислама Каримова!

Только на днях всё прогрессивное человечество отпраздновало Великую дату – 80 - летие Великой Победы в Великой войне над фашизмом! Еще раз поздравляю всех с этой памятной датой и желаю всем доброго здоровья, счастья, благополучия и мира!

Но, наверное не все знают, особенно молодежь, что в марте 1942 года ЛЭТИ был эвакуирован по Ладожской «Дороге жизни» на Большую Землю. План эвакуации предусматривал размещение в Ессентуках, но немецкое наступление продвигалось по стране слишком быстро. 4 августа институт получил приказ о срочной эвакуации и коллектив ЛЭТИ уходил по Кизлярской степи через Каспий, через Туркмению в Ташкент.

25 октября 1942 года ЛЭТИ возобновил занятия на базе Среднеазиатского индустриального института! Сегодня это Ташкентский государственный технический университет!

Коллектив ЛЭТИ активно в этот период помогал становлению науки и образования в Средней Азии и в особенности в нашем университете ТашГТУ! В мае 1944 года была организована первая научно-техническая конференция ППС ЛЭТИ.

В марте 1945 года коллектив института вернулся в родные стены в Ленинград.

Я очень рад доложить сегодня, что эта связь времен и крепкое сотрудничество продолжается до сих пор, сейчас в ТашГТУ находится Проректор ЛЭТИ по международной деятельности Минина Анастасия Андреевна и заместитель генерального директора по персоналу Госкорпорации «Росатом» Терентьева Татьяна Анатольевна в целях подписания Договора о совместной сетевой программе двойного диплома (1+1) «Квантовая фотоника и квантовые технологии». Договор подписан уважаемым ректором ЛЭТИ Шелудько Виктором Николаевичем и ректором ТашГТУ Турабджановым С.М. и первые группы магистрантов по формату 1+1 начнут учебный процесс уже с 2025-26 учебного года!

Уважаемые друзья, участники конференции!

Еще раз повторюсь, считаю большой честью для себя лично приветствовать всех участников этой международной научной конференции, призванной внести свою достойную лепту в вовлечение талантливой молодежи в становление и дальнейшее развитие науки настоящего и будущего, теорию и практику научно-исследовательской деятельности по приоритетным, в настоящее время, и перспективным областям знаний!

Мое приветственное слово закончено и я могу вкратце перейти к основному докладу!

В XXI век искусственный интеллект и индустриальная автоматизация вошли как приоритетные и наиболее перспективные направления развития человеческой цивилизации.

Представляется вполне закономерным и естественным, что входя в нынешнее столетие эти направления потребовали от высшей школы соответствующего кадрового обеспечения. В этом плане наш университет играет не маловажную роль.

В Республике реализуются комплексные меры по активному развитию цифровой экономики, а также широкому внедрению современных информационно-коммуникационных технологий во все отраслях и сферах, основные из них государственное управление, образование, здравоохранение и сельское хозяйство.

Нормативно-правовая база обеспечена полностью: слайд 1, слайд2.

По инициативе и при поддержке Международной Ассоциации по искусственному интеллекту еще в 2000 году было принято решение о проведении один раз в два года в Ташкенте, в стенах нашего Ташкентского государственного технического университета Всемирной научной конференции «Интеллектуальные системы в индустриальной автоматизации». Тем самым международное научное сообщество высоко оценило достойный вклад ученых и специалистов, работающих в области разработки методов искусственного интеллекта в промышленной автоматизации.

Мне отрадно отметить, что из года в год от одного инновационного проекта к другому у нас в Узбекистане реализуются инновационные проекты в топливно-энергетическом комплексе, химической промышленности, в нефте- и газодобывающей промышленности, в которых воплощаются идеи и разработки, доложенные и обсужденные на предыдущих международных конференциях. В списке таких систем индустриальной автоматизации можно назвать: «Навои Азот», «Фергана Азот», «Максам Чирчик», «Максам Аммофос», Шуртанский газохимический комплекс, Кунградский содовый завод, ФНПЗ, Бухарский нефтеперерабатывающий завод, НГМК и др.

Еще раз желаю участникам конференции успешной работы и новых творческих свершений в сфере развития промышленный автоматизации и управления.

Спасибо за внимание!

СЕКЦИЯ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ПРИКЛАДНЫХ ОБЛАСТЯХ

ФОРМИРОВАНИЕ МНОГОМЕРНОГО УПРАВЛЕНИЯ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

БАЗАРБАЕВ К.А., МАННАНОВ У.В.

Ташкентский государственный технический университет имени И.Каримова

Аннотация. Предложена методика синтеза нейронных сетей для идентификации параметров систем по переходным характеристикам. Разработанный подход позволяет эффективно решать задачи синтеза многомерных систем управления, используя современные программные инструменты. Методика демонстрирует высокую точность параметрической идентификации и может применяться для оптимиза ции реальных объектов управления. Результаты исследования открывают новые возможности для развития нейросетевых методов в теории управления.

Ключевые слова: *нейронные сети, идентификация параметров, системы управления, переходные характеристики.*

Введение.

Одними из наиболее сложных и комплексных объектов автоматизации являются многомерные объекты, одной из главных особенностей которых является наличие нескольких регулируемых и регулирующих параметров с тесной взаимосвязью между ними. Подобные взаимосвязи приводят к усложнению регулирования, обусловленному тем, что при регулировании дополнительно требуется учитывать влияние на вторую пару регулирующих параметров, а также влияние второй пары на первую. Причём с ростом количества параметров сложность возрастает кратно.

При этом существует такой набор методов, как искусственные нейронные сети, особенностью которых является возможность обучения решению широкого спектра задач.

Очевидным кажется вариант упрощения синтеза многомерных систем управления при помощи искусственных нейронных сетей, комплекс из которых, обученных для разных задач, будет брать на себя часть работы, упрощая синтез подобной системы.

1.1. Выбор структуры и синтез упрощённой нейронной сети

У искусственных нейронных сетей может быть различная структура, её обычно разделяют по количеству слоёв – на однослойные, с 1 скрытым слоем нейронов, и многослойные – с несколькими слоями, зачастую разбитыми на слои с различными функциями.

Другим способом разделения искусственных нейронных сетей по структуре является разделение по структуре взаимодействия слоёв на:

— прямые (feedforward) – обычно достаточно простые нейронные сети, часто однослойные, хотя встречаются и многослойные, главным плюсом явля-

ется скорость обучения, которая обычно крайне высока, однако ограничения способов обучения требуют огромных объёмов оперативной памяти даже для простых систем;

— рекуррентные – представляющие собой нейронные сети с обратной связью внутри или между слоями, существуют обычно только в вариантах многослойных сетей, из-за специфики структуры обучаются медленнее, но подходят для решения большего спектра задач и менее требовательны к физическим ресурсам.

Мы примем разделение по структуре взаимодействия слоёв.

Поскольку мы определились с определением структуры, пришло время перейти к постановке базовой задачи.

2.1. Постановка задачи

В качестве базовой задачи для синтеза упрощённой нейронной сети выступит следующее:

синтез нейронной сети для определения коэффициента усиления передаточной функции неизвестного порядка в диапазоне со второго по четвёртый порядок по отрезку переходной характеристики фиксированной длины с фиксированным временным шагом.

Для решения данной задачи произведём синтез двух нейронных сетей: одной прямой и одной рекуррентной.

Первым шагом для синтеза нейронной сети является подготовка набора данных для обучения и валидации. В нашем случае условия достаточно просты, в связи с чем мы можем сгенерировать данный набор данных, однако для оптимизации процесса обучения и синтеза введём два дополнительных ограничения:

1. Для создания набора данных будут использоваться только устойчивые функции, разнообразие вариаций устойчивых функций уже достаточно велико для усложнения процесса обучения, однако главным является то, что у неустойчивых передаточных функций невозможно определить коэффициент усиления, а значит они не подходят для обучения.

2. Установим верхний и нижний пределы для значений переходной характеристики в ±25 для валидационных данных и ±35 для данных используемых для обучения. Появление значений больше данных ограничений у реального объекта уже означает достаточно высокую чувствительность по параметру, а подобные системы малопригодны для идентификации методами с точностью менее 99.99%, поскольку могут привести к возникновению опасных ситуаций на предприятии.

Исходя из данных ограничений создадим функцию для проверки устойчивости передаточной функции и функцию генератор.

Функция для проверки устойчивости по критерию Рауса-Гурвица:

Вид и структура полученного набора данных приведены в Приложении 1. Да лее произведём синтез прямой нейронной сети, для чего используется модуль nftool, вид которого показан на рисунках 2.1 и 2.2. На рисунках 2.3. и 2.5 показаны изменения результатов работы сети в процессе обучения, на рисунке 2.4 показана гистограмма отклонений обученной сети, а на рисунке 2.6 – графики соответствия результатов работы сети по данным, используемым для обучения, валидации и тестов.

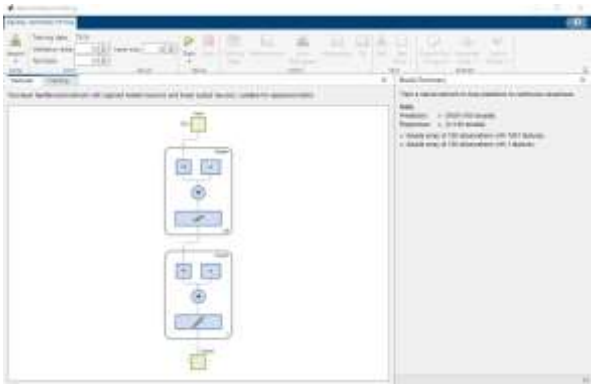


Рисунок 2.1 - Вид Neural network fitting tool

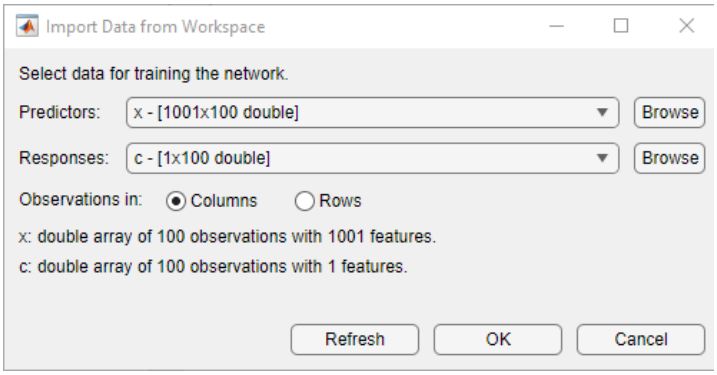


Рисунок 2.2 – Подготовка данных для обучения

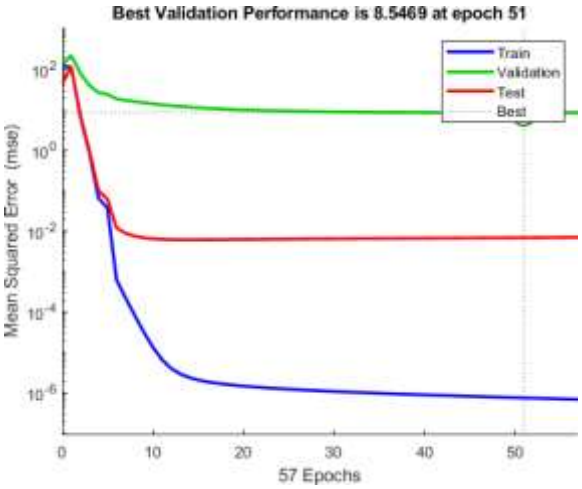


Рис. 2.3 – График изменения отклонения в процессе обучения

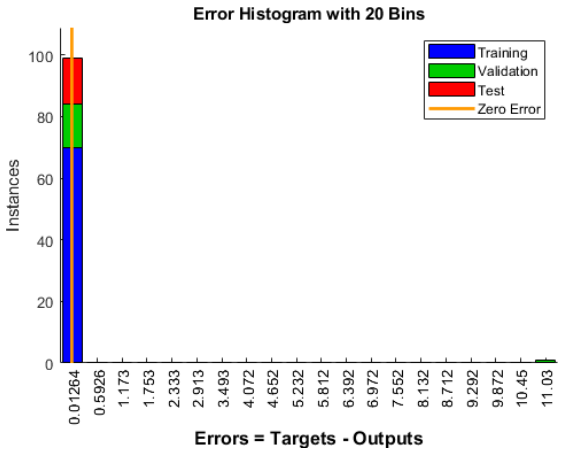


Рис. 2.4 Гистограмма отклонений итоговой сети

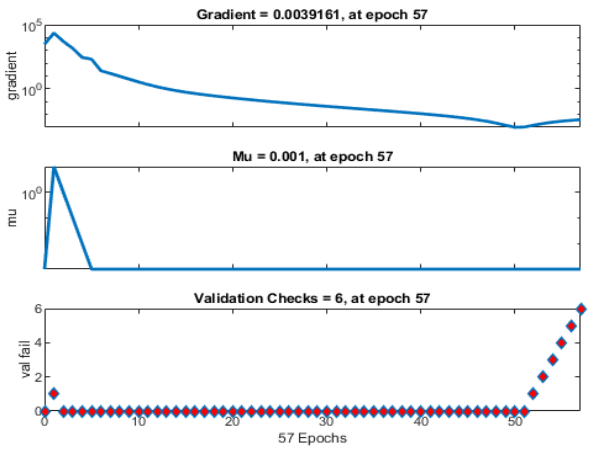


Рис. 2.5 – Графики процесса обучения

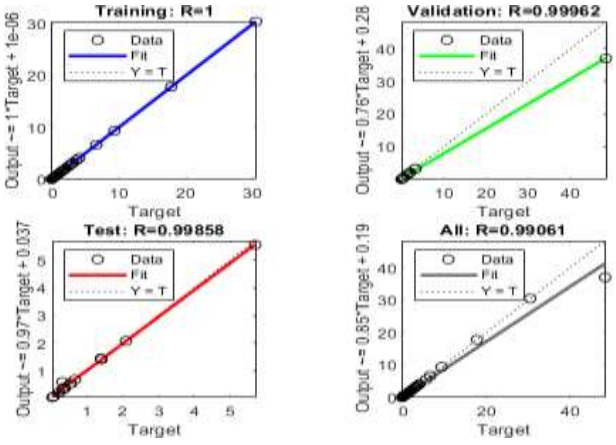


Рис. 2.6 – Графики изменения параметров

Поскольку прямую нейронную сеть мы уже обучили, обучим рекуррентную сеть. Для чего создадим переменные опций обучения и слоёв нейронной сети. В результате,

представленном на рисунке 2.7, где приведены изменения среднего квадратичного отклонения (СКО) и ошибок нейронов в процессе обучения, были получены две нейронные сети с разными структурами.

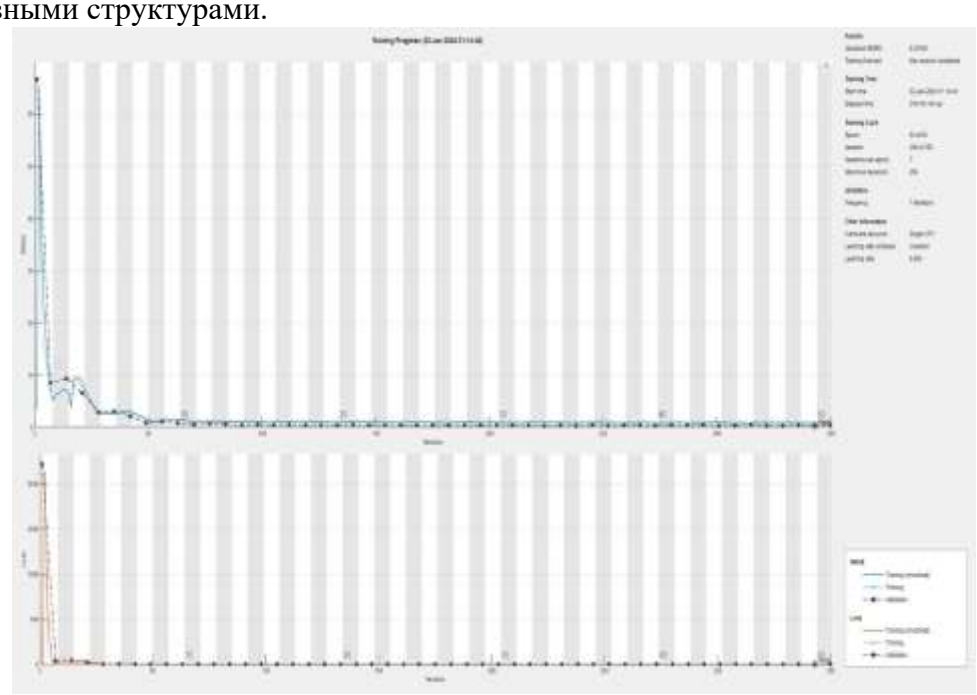


Рис. 2.7 – Результат обучения рекуррентной нейронной сети

Обсуждение полученных результатов.

Для проверки зададим следующие параметры:

Проверка производится при помощи 100 переходных характеристик с требуемым отклонением не более 5%. По итогу точность рекуррентной нейронной сети невысокая, однако краткая сводка по обучению и параметрам данных нейронных сетей следующая:

- Прямая нейронная сеть:

25 нейронов в скрытом слое

Требование ОЗУ для обучения – 22-24 гб

Время обучения – 2,3 часа

- Рекуррентная нейронная сеть:

1000 нейронов для анализа временных взаимосвязей

500 нейронов в полностью соединённом слое

Требование ОЗУ для обучения – 4,6 гб

Время обучения – 5,25 часа

Отсюда видно, что разница в требованиях к оперативной памяти – огромна, исходя из чего можно сделать вывод что синтез сложной прямой нейронной сети не представляется возможным без специализированного оборудования, в частности требованием для сети, способной на вторую поставленную задачу является 497,6 гб оперативной памяти только под её обучение. Это и обуславливает переход на работу с рекуррентными нейронными сетями, хотя на данный момент прямые нейронные сети справляются с выполнением задачи лучше.

Общее заключение о синтезе и структуре нейронной сети

Как видно из данных, приведенных выше, главным минусом прямых нейронных сетей является то, сколько ресурсов требуется на их обучение. При этом их обучение (частично обусловлено их структурой) происходит быстрее. Главными же достоинствами подобных сетей являются малый размер итоговой нейронной сети – файл, содержащий данные полностью обученной сети, занимает лишь несколько мегабайт, а также их простота, что позволяет несколько лучше исправлять возможные проблемы при обучении.

С другой стороны, находятся куда более сложные рекуррентные нейронные сети, сложность их структуры не позволяет просто исправлять ошибки при их обу- чении, также усложнённая структура значительно увеличивает объём, занимаемый файлом с данными нейронной сети, она также усложняет реализацию нейронной сети и её обучение, в связи с чем код для генерации подобной сети более громозд- кий, а само обучение происходит дольше, однако иные алгоритмы, используе- мые в процессе, позволяют значительно сократить требования к ресурсам, за ис- клю- чением требований к вычислительной мощности оборудования.

К сожалению, в нашем случае не получилось показать правильное соотношение точности, поскольку при повторных проверках точность прямой нейронной сети держалась в диапазоне от 68% и вплоть до 75%, в то время как точность рекуррентной нейронной сети не превышала 20%. Однако в среднем, при правильных настройках, рекуррентная нейронная сеть точнее примерно на пять-семь процентов, что на самом деле значительное увеличение.

Список литературы

1. Интернет книга «Neural Networks and Deep Learning» за авторством Mi- chael Nielsen [Электронные ресурс]/Режим доступа: http://neuralnetworksanddeeplearning.com/

2. Официальный сайт «MathWorks» [Электронные ресурс]/Режим доступа: https://uk.mathworks.com/help/deeplearning/ug/sequence-to-sequence- regression-using-deep-learning.html?s_tid=srchtitle_site_search_3_sequence

АНАЛИЗ И ПРОВЕРКА ТЕКСТА НА НАЛИЧИЕ ИСКУССТВЕННО СГЕНЕРИРОВАННЫХ ЧАСТЕЙ С ПОМОЩЬЮ LLM

А.Р Абдулаев

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И.Ульянова (Ленина)

Аннотация В этой статье рассматриваются возможности анализа текста и обнаружение потенциальных участков с присутствием искусственно сгенерированных блоков с помощью большой языковой модели (LLM), состоящей из шести блоков кодировщиков и сверточной нейронной сети (CNN) в качестве классификатора, а также рассматривается задача извлечения ключевых слов из текста используя связи полученные при обучении модели.

Ключевые слова: Анализ, текст, нейронные сети, трансформер, кодировщик, классификация.

Актуальность и востребованность рассматриваемого подхода

На данный момент в области обработки естественного языка существует много различных методов анализа и классификации текста, например ранние методы

классификации авторов произведений с помощью метода “мешка слов”, но такие методы не учитывают контекст и смысл текста. Благодаря сравнительно недавно появившимся трансформерным архитектурам [1] появилось много новых способов для анализа или даже генерации текста с помощью алгоритма, который учитывает связи между всеми словами во входящей последовательности. В этой работе рассматриваются варианты использования такой архитектуры для идентификации искусственно написанного текста и извлечения ключевых слов из текста.

Описание работы алгоритма

Для начала работы с алгоритмом, необходимо поступающую последовательность слов преобразовать в последовательность чисел (токенов). В данной работе используется токенизатор Rubert-tiny[2], состоящий из 29564 возможных токенов.

Пример:
Вводная строка: “Тестовая строка”
Токенизатор: “2, 12824, 13290, 3261, 5406, 13464, 3”
Перевод в текст: “[CLS], Те, ##стов, ##ая, стр, ##ока, [SEP]”

Дополнительные символы [CLS] часто используется для классификации и [SEP] означает конец последовательности.

Так как числа полученные при токенизации сами не несут достаточно информации о слове или части слова, используются векторные представления для каждого возможного токена. Например, при использовании таких векторов расстояние между условными токенами “Автобус” и “Машина” будет значительно меньше, чем между токеном “Машина” и токеном “Бумага”. В данной работе используются представления размером 512 для каждого токена. После этого к векторным представлениям добавляются данные о позиции токена в последовательности. Существует много различных методов её обозначения, в данной работе используется набор параметров, который настраивается при обучении всей модели. Далее следует самый главный блок множественного внимания (Multi-head attention), который считает зависимость между всеми токенами в последовательности. Для этого используется формула 1:

$$Attention(Q, K, V) = soft\max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

Где Q - Запрос, матрица из векторов входной последовательности умноженная на обучаемую матрицу.

K - Ключ, матрица из векторов входной последовательности умноженная на обучаемую матрицу, используется для оценки соотношения токенов с запросом (Q).

V - Значение, матрица из векторов входной последовательности умноженная на обучаемую матрицу, которая умножается на получившеюся матрицу весов внимания.

В данной модели используются 6 идентичных слоев множественного внимания, каждый имеет 8 параллельных голов. Суммарно с базовым классификатором для обучения в модель состоит из 49,283,964 параметров.

Для обучения модели использовался метод скрытия произвольных токенов в последовательности и угадывания изначального значения моделью. При обучении были использованы тексты из Wikipedia (4.7Гб), набор текста из Habr (2.4 Гб), а также из набора

“Taiga” (1.1 GB) Суммарно, обучение длиной в 2 эпохи заняло около 18 часов. Из-за своей длительности процесс обучения был разбит на несколько этапов.

Таблица 1

Размер эмбеддингов (представления)	Макс.длина последовательности	Количество голов	Количество слоев	Начальная скорость обучения	Исключение	Размер пакета (батча)
512	128	8	6	5 x 10^{-5}	0.2	32

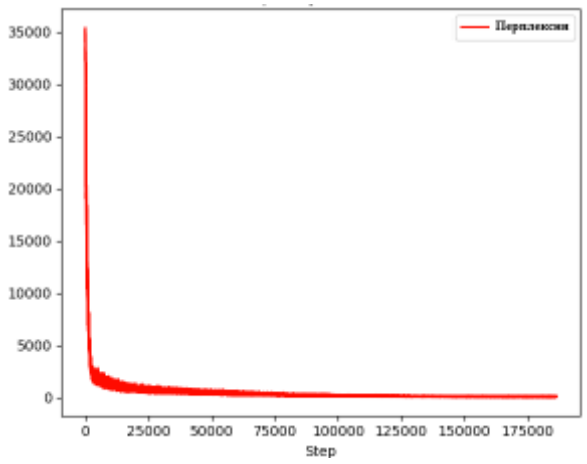


Рис. 2

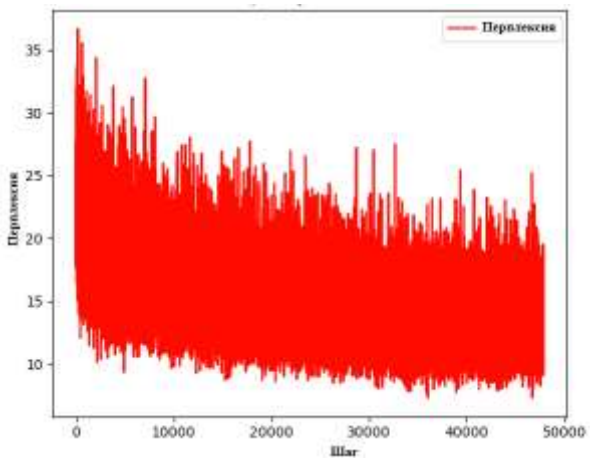


Рис. 3

Можно заметить, что значение перплексии меняется в зависимости от редкости угадываемых токенов. При общем объеме словаря = 29564, если выбор между токенами был бы случайным, то перплексия была бы равна 29564.

Связи, полученные в результате обучения, можно увидеть при извлечении ключевых слов методом вычисления косинусного коэффициента между кодировкой всей последовательности и кодировкой отдельной группы токенов по формулам 2,3 [2]:

$$S_c(A,B) = \cos(\theta) = \frac{A \cdot B}{||A||B||} (2)$$

$$D_c(A,B) = 1 - S_C(A,B) (3)$$

Пример входной последовательности, взятой из данной статьи:

“Так как числа полученные при токенизации сами не несут достаточно информации о слове или части слова, используются векторные представления для каждого возможного токена. Например, при использовании таких векторов расстояние между условными токенами “Автобус” и “Машина” будет значительно меньше, чем между токеном “Машина” и токеном “Бумага”. В данной работе используются представления размеров 512 для каждого токена.”

Слова с самым низким коэффициентом:

1. “токен”, 0.47973281145095825
2. “токенизация”, 0.48263436555862427
3. “вектор”, 0.5455149412155151

Данный алгоритм не только позволяет извлечь самые главные слова из текста, но и демонстрирует успешную работу модели.

Для задачи классификации есть несколько подходов, например похожая по структуре модель “BERT”[4] (трансформерная модель, состоящая только из блоков шифровки от Google) использует начальный токен [CLS] для вывода данных для классификации. В данной модели для классификации используются все выходные данные, полученные в результате кодировки (для одной последовательности размер выходных данных 128x512). Выходные данные используются в свёрточной нейронной сети (CNN).

Для тестирования модели были сгенерированы искусственные абзацы текста при помощи Chat-GPT 4 и GigaChat 2.0 (суммарно 267), и были использованы различные отрывки текста, написанного человеком (суммарно 106), включая, например литературу, отчеты по курсовой работе, и текст из учебника. Были исследованы классификаторы со сверточной сетью и с сетью с прямой связью, а также для сравнения был исследован метод классификации TF-IDF с логистической регрессией. Результаты представлены в таблице 1.

Таблица 1

Метод	Точность (Человеческий текст)	Точность (Искусственный)	Точность (Общая)
LLM+CNN	86.79%	95.88%	93.30%
LLM+FFN	87.7%	92.1%	90.88%
TFIDF + Логистическая регрессия	82.1%	89.88%	87.67%

Заключение

Методы с использованием LLM для классификации текста работают лучше, чем более простые методы подсчета слов, так как они способны учитывать контекст и понимать связи между словами, что помогает при выявлении неочевидных признаков искусственного текста даже с использованием простого кодировщика с относительно малым количеством параметров. Возможно, что при использовании более сложной архитектуры и большего количества текста для обучения точность была бы значительно выше.

Список литературы

1.Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. Attention is all you need //Advances in neural information processing systems. – 2017. – Т. 30.

2. DeepPavlov rubert-base-cased- https://huggingface.co/DeepPavlov/rubert-base-cased

3. COSINE DISTANCE COSINE SIMILARITY ANGULAR COSINE DISTANCE ANGULAR COSINE SIMILARITY [Электронный ресурс]. URL: https://www.itl.nist.gov/div898/software/dataplot/refman2/auxillar/cosdist.htm
4. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova Bert: Pre-training of deep bidirectional transformers for language understanding //Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers). – 2019. – С. 4171-4186.

КЛАСТЕРНЫЙ АНАЛИЗ ПРОФЕССОРСКО-ПРЕПОДАВАТЕЛЬСКОГО СОСТАВА С ИСПОЛЬЗОВАНИЕМ NLP

Будило Е.А.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Аннотация. В работе представлены результаты кластерного анализа преподавателей СПбГЭТУ «ЛЭТИ» на основе данных о научном руководстве, публикационной и преподавательской деятельности. Исследование охватывает сведения о 445 научных руководителях. Для формирования признаков преподавателей использована нейросетевая модель paraphrase-multilingual-mpnet-base-v2. В качестве методов кластеризации применялись алгоритмы K-means, DBSCAN, агломеративная и спектральная кластеризации. Эффективность кластеризации оценивалась по метрике силуэтного коэффициента, отражающего компактность и разделимость кластеров. Наилучшие результаты достигнуты при использовании агломеративной кластеризации (силуэт 0.2129). Полученные результаты демонстрируют потенциал применения векторных представлений и современных методов кластеризации для анализа профессиональной деятельности преподавателей.

Ключевые слова: кластеризация, обработка естественного языка, машинное обучение, анализ образовательных данных, текстовые эмбеддинги

Введение

Современные исследования подчеркивают активное внедрение ML-технологии в высшем образовании для решения задач прогнозирования, оценки и персонализации образовательных процессов, что способствует формированию новых исследовательских тенденций и росту публикационной активности в этой области [1]. Одним из перспективных направлений является применение глубоких нейронных сетей для комплексной и объективной оценки качества преподавания, что позволяет повысить точность анализа и минимизировать влияние субъективных факторов [2].

Цель работы — кластеризация преподавателей СПбГЭТУ «ЛЭТИ» по направлениям деятельности на основе векторных признаков, сформированных из данных о научном руководстве, публикационной и преподавательской деятельности с использованием нейросетевой модели paraphrase-multilingual-mpnet-base-v2 [3]. Для этого поставлены задачи: формирование признаков по трём категориям данных, применение и сравнительный анализ алгоритмов кластеризации, а также интерпретация выявленных кластеров.

Описание и предобработка данных

В исследовании использована база данных, включающая сведения о 445 преподавателях СПбГЭТУ «ЛЭТИ», курировавших выпускные квалификационные работы (ВКР) в 2024 году. Для каждого преподавателя зафиксированы количество (от 1 до 44, среднее – 5.074) и тематика ВКР, а также темы аспирантских диссертаций, что позволяет комплексно охарактеризовать направления научного руководства.

Кроме того, в наборе данных содержится научно-исследовательская активность преподавателей и метаданные публикаций (название, DOI, ISBN, индексация в WoS, Scopus, РИНЦ и др.), а также ключевые слова, аннотации и тематические категории при их наличии. Набор данных дополняется сведениями о патентах и регистрациях программ для ЭВМ (РИД) с указанием названия и области техники. Важной частью данных являются рабочие программы дисциплин, содержащие подробное описание содержания курсов, перечень тем и экзаменационных вопросов, используемые для анализа структуры и содержания образовательного процесса.

Для формирования векторных признаков использована мультиязычная нейросетевая модель paraphrase-multilingual-mpnet-base-v2, преобразующая текстовые данные (русский, английский и др.) в 1024-мерные эмбеддинги, сохраняющие семантические и синтаксические особенности текста.

Векторные признаки формировались на основе трёх категорий данных: научного руководства (темы ВКР и диссертаций), научно-исследовательской деятельности (публикации с аннотациями и РИД), а также преподавательской деятельности (рабочие программы дисциплин). Для публикаций входной строкой в модель подавалось объединение названия и аннотации, что повышало информативность получаемых векторов.

Итоговое векторное представление каждого преподавателя формировалось с помощью взвешенного усреднения эмбеддингов: коэффициенты составляли 0.3 для публикаций и дисциплин, 0.2 для тем ВКР и 0.1 для тем диссертаций и РИД, что позволило минимизировать влияние шумовых данных и сохранить релевантность итогового признакового пространства для последующего кластерного анализа.

Кластеризация и оценка качества

Для решения задачи кластеризации применялись алгоритмы K-means [4], DBSCAN [5], агломеративная [6] и спектральная кластеризации [7].

K-means минимизирует внутрикластерную сумму квадратов расстояний, итеративно обновляя центроиды до достижения сходимости. DBSCAN выделяет кластеры как области высокой плотности, используя параметры eps и min_samples, и автоматически определяет их количество, включая обработку выбросов. Агломеративная кластеризация формирует иерархию объединением ближайших кластеров на основе мер близости, завершая процесс при достижении целевого числа кластеров. Спектральный метод преобразует данные через спектральный анализ матрицы сходств, снижая размерность перед применением классических алгоритмов, что позволяет выявлять кластеры сложных форм. Все методы демонстрируют уникальные преимущества в зависимости от структуры данных.

Для анализа эффективности кластеризации применялась метрика силуэтного коэффициента (далее – силуэта) [8], количественно характеризующая компактность кластеров и степень их взаимной отделимости. Значение силуэта для объекта определяется как:

$$Sil(C) = \frac{1}{N} \sum_{c_k \in C} \sum_{x_i \in c_k} \frac{b(x_i, c_k) - a(x_i, c_k)}{\max\{a(x_i, c_k), b(x_i, c_k)\}},$$

где $a(x_i, c_k) = \frac{1}{|c_k|} \sum_{x_j \in c_k} \|x_i - x_j\|$ – среднее расстояние от $x_i \in c_k$ до других объектов из кластера c_k (компактность), $b(x_i, c_k) = \min_{c_l \in C \setminus c_k} \left\{ \frac{1}{|c_l|} \sum_{x_j \in c_l} \|x_i - x_j\| \right\}$ – среднее расстояние от $x_i \in c_k$ до объектов из другого кластера $c_l : k \neq l$ (отделимость).

Метрика силуэта принимает значения в диапазоне $-1 \leq Sil(C) \leq 1$, где значения, близкие к 1, указывают на оптимальное разделение кластеров, а отрицательные значения сигнализируют о потенциальной ошибочной кластеризации. Чем выше значение $Sil(C)$, тем выше внутрикластерная схожесть и межкластерная различимость.

Результаты

Все эксперименты воспроизводились с помощью библиотеки scikit-learn [9] для Python. Для алгоритмов агломеративной кластеризации, спектральной кластеризации и K-means была выполнена серия экспериментов с перебором числа кластеров. Зависимость значения метрики силуэта от количества кластеров представлена на рисунке 1.

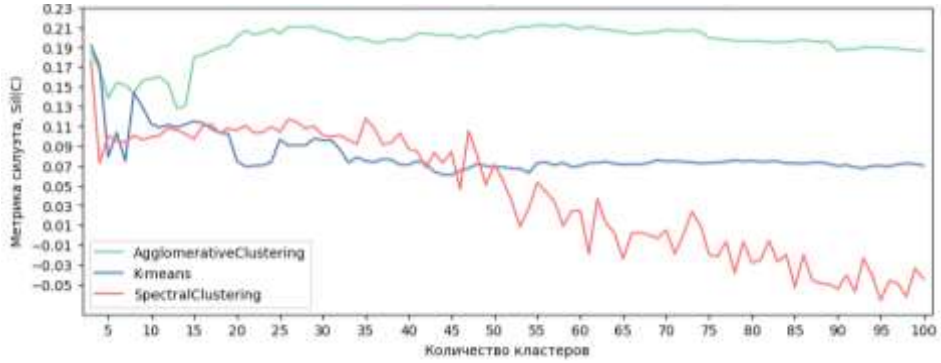


Рис. 1. Зависимость метрики силуэта от количества кластеров

В случае спектральной кластеризации исследовался параметр assign_labels, определяющий стратегию присвоения номеров кластеров в пространстве признаков. Наилучший результат был достигнут при значении assign_labels = "kmeans" (0.1180), что превосходит показатели при использовании стратегий "discretize" (0.0976) и "cluster_qr" (0.1078).

Для агломеративной кластеризации варьировался параметр linkage, отвечающий за способ объединения кластеров. Оптимальным оказалось значение linkage = "average" с метрикой силуэта 0.2129, что значительно выше по сравнению с "single" (0.0859) и "complete" (0.1796).

В алгоритме DBSCAN проводился перебор параметров eps в диапазоне от 0.001 до 0.1 и min_samples от 2 до 5. Лучшие результаты получены при eps = 0.02 и min_samples = 4.

Сводные результаты по каждому методу кластеризации представлены в таблице 1.

Таблица 1

Метрика силуэта при разных подходах кластеризации

№	Алгоритм кластеризации	Количество кластеров	Параметры	$Sil(C)$
1	Агломеративная кластеризация	58	metric = "cosine", linkage = "average"	**0.2129**
2	K-means	8	—	0.1440
3	DBSCAN	5	eps = 0.02, min_samples = 4	0.0556
4	Спектральная кластеризация	35	assign_labels = "kmeans"	0.1180

Наилучший результат по метрике силуэта показала агломеративная кластеризация (0.2129). Визуализация результатов методом t-SNE [10] представлена на рисунке 2.

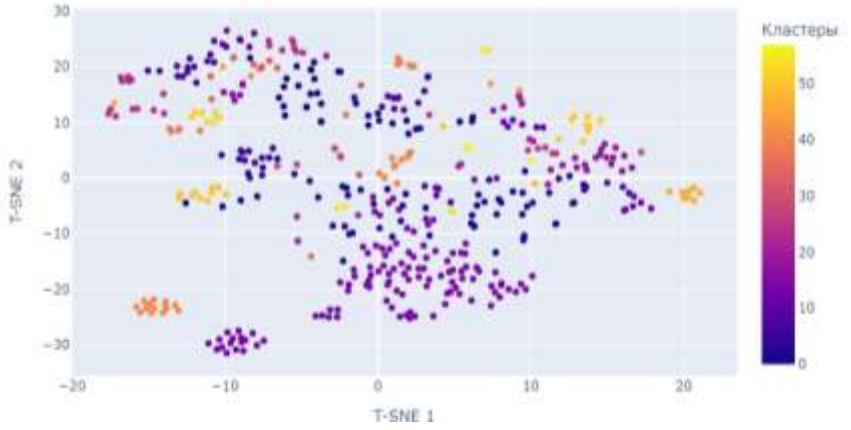


Рис. 2. Визуализация результатов агломеративной кластеризации в двумерном пространстве признаков

На рисунке 2 отчетливо выделяются два крупных кластера в нижней левой части: оранжевым обозначен кластер преподавателей, связанных с маркетингом и коммуникационными продвижениями, фиолетовым – с лингвистикой. Жёлтый кластер в верхней части соответствует экологическим наукам.

Заключение

В ходе исследования был проведён кластерный анализ профессорско-преподавательского состава СПбГЭТУ «ЛЭТИ» на основе интеграции данных о научном руководстве, публикационной и преподавательской деятельности с применением современных методов обработки естественного языка и нейросетевых векторных представлений. Сравнительный анализ алгоритмов кластеризации показал, что агломеративная кластеризация обеспечивает наилучшее качество разделения преподавателей по профессиональным направлениям (максимальный силуэтный коэффициент 0.2129), что подтверждает релевантность выбранного признакового пространства и методологии обработки данных.

Полученные кластеры отражают реальные различия в профилях деятельности (например, выделены группы по маркетингу, лингвистике, экологическим наукам), что открывает возможности для объективной сегментации кадрового состава, оптимизации кадровой политики и индивидуализации образовательных траекторий. Результаты работы демонстрируют высокую значимость применения векторных эмбеддингов и машинного обучения для анализа образовательных данных и могут служить основой для создания рекомендательных систем и поддержки управленческих решений в вузах. Перспективы дальнейших исследований связаны с расширением спектра анализируемых данных, валидацией моделей и внедрением разработанных подходов в практику управления образовательными организациями. Реализация этих направлений позволит повысить конкурентоспособность вуза и сформировать научно обоснованные подходы к развитию кадрового потенциала в высшем образовании.

Список литературы

1. Tahiru F., Parbanath S., Agbesi S. Machine learning-based predictive systems in higher education: a bibliometric analysis //Journal of Scientometric Research. – 2023. – Т. 12. – №. 2. – С. 436-447.

2. Liu Y. Evaluation Algorithm of Teaching Work Quality in Colleges and Universities Based on Deep Denoising Autoencoder Network //Mobile Information Systems. – 2021. – Т. 2021. – №. 1. – С. 8161985.
3. Reimers N., Gurevych I. Sentence-bert: Sentence embeddings using siamese bert-networks //arXiv preprint arXiv:1908.10084. – 2019.
4. MacQueen J. Some methods for classification and analysis of multivariate observations //Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics. – University of California press, 1967. – Т. 5. – С. 281-298.
5. Ester M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise //kdd. – 1996. – Т. 96. – №. 34. – С. 226-231.
6. Ward Jr J. H. Hierarchical grouping to optimize an objective function //Journal of the American statistical association. – 1963. – Т. 58. – №. 301. – С. 236-244.
7. Ng A., Jordan M., Weiss Y. On spectral clustering: Analysis and an algorithm //Advances in neural information processing systems. – 2001. – Т. 14.
8. Rousseeuw P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis //Journal of computational and applied mathematics. – 1987. – Т. 20. – С. 53-65.
9. Pedregosa F. et al. Scikit-learn: Machine learning in Python //the Journal of machine Learning research. – 2011. – Т. 12. – С. 2825-2830.
10. Van der Maaten L., Hinton G. Visualizing data using t-SNE //Journal of machine learning research. – 2008. – Т. 9. – №. 11.

ОБЗОР МОДЕЛЕЙ АВТОМАТИЧЕСКОГО РАСПОЗНАНИЯ АКЦЕНТИРОВАННОЙ РЕЧИ

Волкова А.А., Дружинская Е.В.

Уфимский государственный нефтяной технический университет

Аннотация. В статье рассматриваются особенности акцентированной русской речи у иностранных студентов, обусловленные влиянием родного языка на произносительную сторону речи. Описаны типичные трудности: замены звуков, смешение твёрдых и мягких согласных, нарушение ударения и ритма. Представлен обзор моделей распознавания речи, включая традиционные вероятностные подходы, нейросетевые архитектуры, деревья решений, а также гибридные решения. Рассматриваются их особенности и применимость для обработки акцентированной речи в образовательных целях. Полученные результаты могут быть использованы при разработке адаптивных речевых интерфейсов и образовательных платформ, ориентированных на анализ неродной речи и сопровождение процесса обучения произношению.

Ключевые слова: автоматические системы, распознавание речи, акцентированная речь, нейронные сети, модели машинного обучения, акцент.

Современные системы автоматического распознавания речи демонстрируют высокую точность при работе с речью носителей языка, однако при обработке акцентированной речи их эффективность снижается. Это ограничивает применение таких технологий в обучающих приложениях для изучения русского языка как иностранного.

Основные сложности связаны с фонетической и интонационной вариативностью: изменениями в звуковом составе, артикуляции, ударении и ритмике. В результате увеличивается количество ошибок распознавания, что затрудняет точную интерпретацию речи студентов [1].

Для распознавания акцентированной речи используются современные нейросетевые подходы, включая адаптированные скрытые марковские модели, рекуррентные нейронные сети с механизмами долгосрочной памяти (LSTM) и глубокие сверточные нейросети. Эти модели обеспечивают многослойный анализ речевого сигнала, сочетая акустическую

информацию с языковым контекстом. Их применение возможно при наличии обучающих данных, охватывающих разнообразные акцентные варианты, а также при использовании алгоритмов, способных адаптироваться к индивидуальным особенностям речи конкретного пользователя.

Акцент формируется под влиянием родного языка обучающегося и может проявляться на разных уровнях: от артикуляции отдельных звуков до интонационно-ритмического оформления высказывания. Например, у носителей тюркских языков часто возникают трудности с произношением звуков, отсутствующих в их родной системе ([ы], [щ], [ш], [ж]), а также с различением твёрдых и мягких согласных [2]. Студенты, говорящие на европейских языках, испытывают сложности с подвижным ударением, русской интонацией и артикуляцией некоторых звуков ([р], [л], [х]) [3].

Типичные проявления акцентированной русской речи:

– замены звуков, отсутствующих в родном языке (например, [ы] → [и]);

– смешение твёрдых и мягких согласных;

– ошибки при произнесении слов с разделительными знаками;

– перенос интонационных моделей родного языка;

– нарушение ударения в слове;

– искажение ритма и просодических элементов.

Акцентированная речь требует особого подхода при обучении: важно не только корректировать ошибки, но и понимать, откуда они берутся. Для этого необходимо учитывать фонетическую систему родного языка студента. Автоматизированные системы распознавания речи, способные обрабатывать акцентированную русскую речь, могут помочь студенту самостоятельно отслеживать прогресс, видеть неточности и получать индивидуальные рекомендации по улучшению произношения. Это особенно ценно в образовательной среде, где необходимо учитывать широкий спектр фонетических вариантов.

Сложность обработки такой речи требует использования разных типов моделей, способных учитывать фонетическую изменчивость и адаптироваться к индивидуальным особенностям речи. Современные системы распознавания речи основываются на нескольких подходах, включая вероятностные методы и нейросетевые архитектуры. К вероятностным моделям относится скрытая марковская модель, в которой речь представляется как последовательность фонем, моделируемых через цепи состояний с заданными вероятностями переходов. Она активно применяется для описания временной структуры акустического сигнала, в том числе в задачах, связанных с акцентной речью. Также применяются деревья решений — интерпретируемые модели, формирующие иерархию правил, однако они менее эффективны при работе с большими объёмами аудиоданных.

Среди нейросетевых методов можно выделить глубокие нейронные сети, позволяющие выявлять сложные зависимости в данных за счёт многослойной структуры, сверточные нейросети, применяемые для обработки спектрограмм и других акустических признаков, и рекуррентные сети (RNN), ориентированные на последовательную природу речевого сигнала. Последние особенно полезны при моделировании временного контекста, а их модификации, например, двунаправленные RNN и архитектуры с механизмом долгосрочной памяти, применяются при распознавании акцентной речи, где важно учитывать как предшествующие, так и последующие элементы сигнала. В случае необходимости обработки сложных связей между признаками речи, особенно в

многоуровневом контексте, могут использоваться графовые нейронные сети, способные представлять речевые данные в виде структурированных графов [4].

Для наглядного сопоставления методов в таблице 1 ниже приведены краткие характеристики рассмотренных моделей:

Таблица 1

Особенности моделей распознавания речи с учетом акцентов

Модель	Описание	Преимущества	Недостатки	Особенности при работе с акцентами
Скрытые марковские модели	Статистическая модель переходов между звуками, основанная на вероятностях	Простота, эффективность для небольших моделей	Ограничена в обработке сложных зависимостей	Устойчивы к шуму, но ограниченно адаптируются к вариативности акцентированной речи
Деревья решений	Иерархическая модель, формирующая правила на основе признаков	Интерпретируемость, быстрота обучения	Низкая обобщающая способность на больших и шумных данных	Слабо справляются с вариативностью произношения, подходят для предварительного анализа
Глубокие нейронные сети	Многослойная архитектура для выявления скрытых признаков	Хорошо обучаются на больших объемах данных, выявляют сложные закономерности	Могут переобучаться, требуют большого количества данных	Эффективны при наличии обучающих выборок с акцентами
Сверточные сети	Выделение локальных признаков из спектрограмм или других входных данных	Устойчивы к шуму, эффективны при анализе спектрограмм	Слабо учитывают временную динамику	Подходят для выделения фонетических особенностей акцентированной речи
Рекуррентные сети (RNN, LSTM)	Учет временных зависимостей в последовательностях	Хорошо моделируют контекст, особенно в длительных речевых отрезках	Возможны трудности с обучением на длинных последовательностях	Могут учитывать изменения темпа и ритма, характерные для акцентной речи
Графовые нейронные сети	Представление речи как графа и анализ взаимосвязей между элементами	Моделируют взаимосвязи между признаками,	Сложность реализации, высокая	Требуют сложной подготовки входных данных, находятся

Модель	Описание	Преимущества	Недостатки	Особенности при работе с акцентами
		хорошо работают с нестандартными структурами	вычислительная нагрузка	на стадии активного развития

В рамках задач обучения русскому языку как иностранному, особенно с акцентом на самостоятельную отработку произношения, предпочтение часто отдается нейросетевым архитектурам, способным адаптироваться под индивидуальные речевые особенности студентов. Использование рекуррентных и сверточных нейросетей, а также методов адаптации моделей под конкретные акценты, позволяет повысить точность распознавания и обеспечить обратную связь, направленную на устранение фонетических ошибок.

В заключение, интеграция современных моделей распознавания речи в образовательные приложения открывает возможности для персонализированного подхода к обучению русскому языку как иностранному. Наиболее перспективным решением в условиях акцентированной речи являются рекуррентные нейронные сети с механизмом долгосрочной памяти (LSTM), так как они эффективно учитывают временную структуру речи и позволяют адаптироваться к индивидуальным особенностям произношения. В сочетании с языковыми моделями они обеспечивают высокую точность распознавания и устойчивость к вариативности интонации и акцента. Это особенно актуально в многоязычной образовательной среде, где необходимо учитывать широкий спектр фонетических вариаций и речевых характеристик студентов.

Список литературы

1. Бернштейн С.И. Вопросы обучения произношению применительно к обучению русскому языку иностранцев // Вопросы фонетики и обучение произношению. М.: Изд-во Московского ун-та, 1975.
2. Сравнительная грамматика русского и туркменского языков. Фонетика и морфология / под ред. Н. А. Баскакова. - Ашхабад : Туркмен. изд-во, 1974. - 380 с.
3. Русский акцент в английском: основные особенности и как от него избавиться // Хабр [Электронный ресурс]. – 2021. – URL: https://habr.com/ru/companies/englishdom/articles/538508/ (дата обращения: 10.04.2025)
4. Кипяткова И.С., Карпов А.А., Кулешов С.В., Зайцева А.А. Методы и модели автоматического распознавания речи: Учеб. пособие /СПб ФИЦ РАН. СПб., 2021. 116с.

РАЗРАБОТКА КОМПАКТНЫХ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ДЛЯ ЛОКАЛЬНОЙ ОБРАБОТКИ ДАННЫХ В УСТРОЙСТВАХ ИНТЕРНЕТА ВЕЩЕЙ (IOT)

ГОРБАЧЁВ М.Е.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И.Ульянова (Ленина)

Аннотация. В докладе рассматриваются подходы к созданию компактных нейросетевых моделей, предназначенных для локальной обработки данных на устройствах IoT (интернета вещей). Приводятся методы оптимизации, примеры внедрения в реальных производственных условиях, а также описание разработки и внедрения прототипа.

Ключевые слова: IoT, нейросети, квантование, прореживание, edge computing, оптимизация, машинное обучение

Введение

С развитием технологий IoT (интернета вещей) значительно увеличилось количество подключаемых устройств, что породило необходимость локальной обработки данных прямо на устройствах без отправки в облачные сервисы или на внешние девайсы. Эта потребность особенно важна для промышленных предприятий, интеллектуальных систем в жилых зданиях, носимых медицинских гаджетов и автоматизированного сельского хозяйства. Сегодня компактные нейросетевые модели обеспечивают возможность оперативной и надежной обработки информации, повышая автономность и безопасность устройств.

Постановка задачи

Облачная обработка данных не всегда оптимальна в IoT по следующим причинам:

- Высокие задержки передачи данных
- Угрозы конфиденциальности информации
- Перегруженность каналов связи
- Высокое энергопотребление устройств
- Ограниченность аппаратных ресурсов IoT-устройств

Разработка и внедрение компактных моделей нейросетей способствует решению этих проблем, повышает эффективность систем и минимизирует затраты ресурсов.

Методы оптимизации нейросетевых моделей

Рассматриваются три ключевых метода:

- Квантование - uniform/non-uniform, post-training и quantization-aware training.
- Прореживание - magnitude-based, structured и dynamic pruning.
- Дистилляция знаний - модели типа teacher-student и soft labels.

Применение в различных сферах

1. Промышленность - анализ вибраций, контроль качества, превентивное обслуживание.
2. Умные здания - видеонаблюдение, микроклимат, энергоменеджмент.
3. Здравоохранение - мониторинг, аномалии.
4. Сельское хозяйство - диагностика растений, автоматизация.
5. Носимые устройства - часы, браслеты, распознавание речи.

Примеры успешного внедрения

- Bosch Rexroth - обнаружение дефектов.
- ABB - предсказание поломок.
- Siemens - экономия энергии.
- Garmin - отслеживание здоровья.

Описание реализации компактной модели

Архитектура - MobileNetV3-Small. Обучение - квантование, прореживание. Тестирование на ESP32 ввиду высокой скорости и малая нагрузка.

Список литературы

1. Han S., Mao H., Dally W. Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding. ICLR, 2016.
2. Howard A., et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, 2017.
3. Hinton G., Vinyals O., Dean J. Distilling the Knowledge in a Neural Network, 2015.
4. TensorFlow Lite | ML for Mobile and Edge Devices [Электронный ресурс]: URL: https://www.tensorflow.org/lite (дата обращения: 10.05.2025).

5. Banbury C., et al. Benchmarking TinyML Systems: Challenges and Direction, 2020.
6. Singh D., et al. IoT and machine learning based edge computing for intelligent manufacturing systems, 2021.
7. Zhang C., et al. A Survey on Neural Network Acceleration, IEEE Transactions, 2020.

СЕГМЕНТАЦИЯ ГИСТОЛОГИЧЕСКИХ ИЗОБРАЖЕНИЙ СОСУДОВ ЛЕГКИХ НА ОСНОВЕ ГЛУБОКОГО ОБУЧЕНИЯ

Жук Ф.П.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Исследование направлено на разработку нейросетевой модели для автоматической сегментации сосудов легочной циркуляции на основе КТ-изображений. В работе использован датасет, содержащий срезы с выделенными сосудистыми структурами. В качестве основы предложена адаптированная архитектура ResNet50, модифицированная для задач семантической сегментации за счет использования трансферного обучения. Для повышения точности применены специфические преобразования данных, включая нормализацию и градацию масок. Предложенное решение продемонстрировало высокую точность: Pixel accuracy (0.92) и IoU (0.93). Результаты подтверждают возможность применения алгоритма для стандартизации диагностики легочных патологий, включая гипертензию.

Ключевые слова: нейронные сети, машинное обучение, ResNet50, сегментация, сегментация сосудов, гистологический анализ.

Введение

Заболеваемость легочными сосудистыми патологиями неуклонно увеличивается. В связи с этим гистологическое исследование сосудов малого круга кровообращения играет критическую роль как в доклинических исследованиях, так и в клинической практике. Традиционные методы анализа сопряжены значительными временными затратами и требованием специальной подготовки. Целью работы стала - разработка высокоточных и автоматизированного инструмента диагностики для сегментации сосудов малого круга кровообращения. В последние годы методы глубокого обучения активно применяются для сегментации биомедицинских изображений. Архитектура U-Net [1], демонстрирует высокую точность в выделении клеточных структур за счет skip-connections, сохраняющих пространственные детали. Но большинство исследований сосредоточено на опухолевых тканях, тогда как сегментация сосудов малого круга кровообращения остается менее изученной. В работе «All answers are in the images: A review of deep learning for cerebrovascular segmentation» [2] авторы продемонстрировали потенциал для анализа сосудистых сетей.

Материалы и методы

В рамках исследования для обучения был выбран набор гистологических изображений на kaggle. Набор данных «Microscopy Image Dataset for Quantitative Assessment of Pulmonary Vascular Changes» [3], включает в себя 609 гистологических изображений легких мышей и масок к ним (рис. 1).

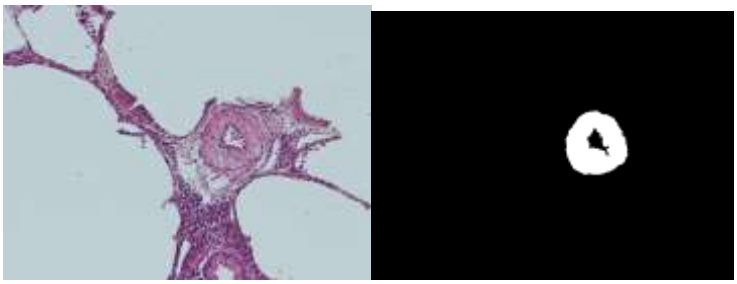


Рис. 1. Пример данных

Данные предобработанны с помощью transforms из библиотеки torchvision. Изображения преобразовались по следующему алгоритму:

1. Исходные изображения в формате PIL конвертируются в тензоры PyTorch, изображение представлено в виде тензора с размерностью [каналы, высота, ширина].
2. Изображения масштабируются до фиксированного разрешения 1280×720, для стандартизация входных данных.
3. Тензоры приводятся к типу torch.float, так как операции глубокого обучения требуют данных в формате с плавающей точкой.
4. Значения пикселей нормализуются по всем каналам с параметрами: mean=[0.5, 0.5,0.5], std=[0.2, 0.2,0.2].

Маски проходят аналогичные шаги 1–4, но с дополнительным преобразованием, оно преобразует три исходных канала изображения в один черно-белый канал, для упрощения задачи сегментации до бинарной классификации.

В качестве модели обучения была выбрана архитектура ResNet50 [4]. Модель настроена следующим образом: количество выходных слоев ровнялось 1, а в качестве исходных весов была выбрана ‘torchvision.models.ResNet50_Weights.IMAGENET1K_V2’ [4].

В качестве метрик для оценки модели использовались Pixel accuracy и IoU accuracy. Метрика Pixel accuracy рассчитывается по следующей формуле:

$$PixelAccuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Метрика IoU accuracy рассчитывается по следующей формуле:

$$IoU = \frac{TP}{TP + FP + FN}$$

,где TP – это количество истинноположительных пикселей, TP – это количество истиннонегативных пикселей, FP – количество ложноположительных пикселей, FN – количество ложноотрицательных пикселей.

ЭВМ, выполняющая задачу обучения, была оснащена ГПУ серии NVIDIA A100 и обладала 30 GB RAM.

Для обучения модели датасет был разбит на 3 набора: 487 экземпляров для обучения, 61 экземпляр для валидации, 61 экземпляр для теста. В качестве оптимизатора был выбран Adam [5] с learning rate = 1e-4, и функция ошибки MSELoss. MSELoss рассчитывается по следующей формуле:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Результаты

В ходе обучения были получены следующие данные продемонстрированные на рисунках 2 и 3.

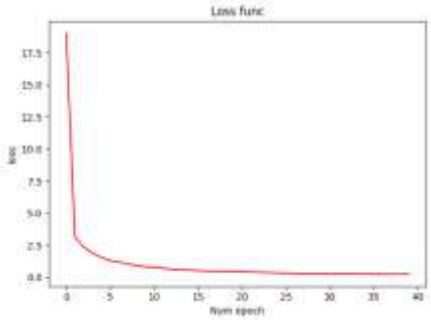


Рис. 2. Динамика функции потерь в ходе обучения

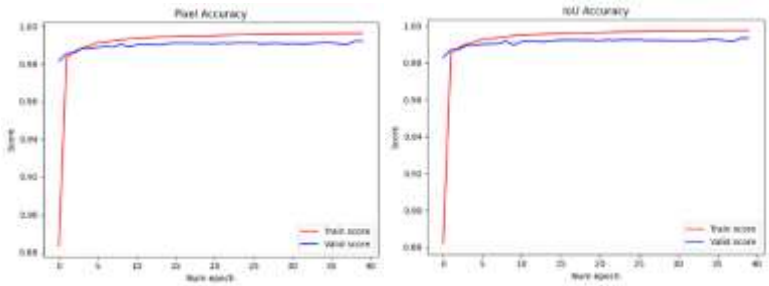


Рис. 3. Динамика метрик в ходе обучения

Модель показала высокое качество на валидационой выборке, а именно выше 0,9 по обеим использованным метрикам качества, что является хорошим результатом.

На тестовой выборке модель показала следующие результаты, продемонстрированные на таблице 1 и рисунке 4.

Таблица 1

Результаты тестовой выборки

Метрика	Среднее значение
Pixel Accuracy	0.92
IoU	0.93
Время обработки 1 изображения	0.18 секунды

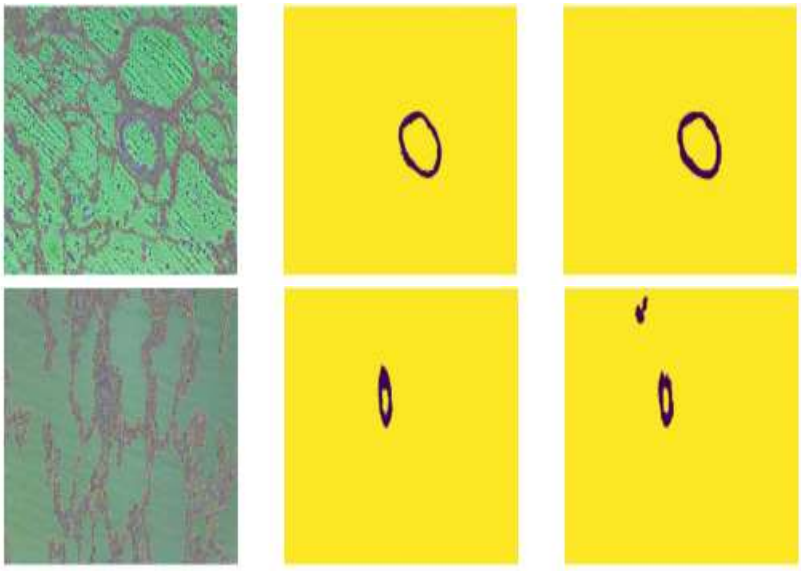


Рис. 4. Сравнение исходных изображений, масок и результатов сегментации

где input image – входное изображение, lable image – эталонное изображение маски, output image – результат работы модели.

Заключение

Практическая значимость работы заключается в создании инструмента для стандартизации гистопатологических исследований, что особенно актуально для диагностики легочной гипертензии и оценки эффективности терапии.

Модель корректно сегментирует сосуды, включая мелкие структуры. Наблюдаются незначительные, такие как шум на границах сосудов, не влияющие на общую точность. В остальном она имеет высокие показатели точности. Кроме того, модель обеспечивает достаточную скорость обработки, что позволяет работать с большими объемами данных в сжатые сроки. На примерах видно, что она может обнаружить даже трудноразличимые для глаза элементы.

Использование единственного датасета, содержащего изображения легких мышей, ограничивает обобщаемость модели для человеческих тканей. Отсутствие полнокадровых гистологических снимков, может снижать точность при работе с реальными клиническими данными. Однако предложенная модель показала эффективность в задачах автоматизированного анализа гистологических изображений сосудов малого круга кровообращения.

Дальнейшие исследования целесообразно направить на создание открытых датасетов с наборами редких патологий сосудов, сокращение количества артефактов, адаптация модели для работы с полнокадровыми гистологическими изображениями.

Предложенный подход подтверждает потенциал глубокого обучения в задачах анализа гистологических данных, однако его внедрение в клиническую практику требует дальнейшего обучения на разнообразных и репрезентативных данных.

Список литературы

1. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // 2015. DOI: 10.48550/arXiv.1505.04597

2. Chen C., Kangneng Z., Zhiliang W., Qian Z., Ruoxiu X. All answers are in the images: A review of deep learning for cerebrovascular segmentation // Computerized Medical Imaging and Graphics. 2023. Т. 107. С. 102229. DOI: 10.1016/j.compmedimag.2023.102229
3. Aleksandr M. Sinitca, Asya I. Lyanova, Dmitrii I. Kaplun, Hassan Hassan, Alexander S. Krasichkov, Kseniia E. Sanarova, Leonid A. Shilenko, Elizaveta E. Sidorova, Anna A. Akhmetova, Dariya D. Vaulina Andrei A. Karpov. Microscopy Image Dataset for Deep Learning-Based Quantitative Assessment of Pulmonary Vascular Changes // Sci Data. 2024. Т. 11. № 1. DOI: 10.1038/s41597-024-03473-z
4. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // 2015. DOI : 10.48550/arXiv.1512.03385
5. Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization // 2014. DOI: 10.48550/arXiv.1412.6980
6. Sudianjaya N. R., Fatichah C. Segmentation and Classification of Breast Cancer Histopathological Image Utilizing U-Net and Transfer Learning ResNet50 // matrik. 2024. Т. 24. № 1. С. 155–166. DOI: 10.30812/matrik.v24i1.4186

ПОСТРОЕНИЕ ТРАЕКТОРИИ ПЕРЕМЕЩЕНИЯ ОБЪЕКТА НА ОСНОВЕ АНАЛИЗА ВИДЕОПОТОКА

ЗДЕХ К.В, ВИКСНИН И.И.

Санкт-Петербургский Государственный Электротехнический Университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В работе решается задача восстановления глобального движения сцены по видеопотоку без геолокации. Предложен метод на основе алгоритма LoFTR для извлечения плотных соответствий и последующего предсказания аффинных преобразований (угол поворота и сдвиг) с помощью нейронной сети. Метод сравнивался с классическими подходами (Лукас-Канаде, Фарнебак, сверточные сети) и показал более высокую точность и устойчивость при больших смещениях, размытии и изменениях разрешения. Результаты подтверждают эффективность сочетания LoFTR и нейронной регрессии для анализа визуальных траекторий.

Ключевые слова: оптический поток, LoFTR, аффинное преобразование, анализ движения, нейронные сети, регрессия, сопоставление признаков, визуальная одометрия, трансформеры, обработка изображений

Введение

В современных задачах анализа перемещений объектов часто возникает необходимость определения маршрутов движения при отсутствии доступа к данным геолокации. В таких случаях альтернативным решением становится использование видеопотока с камер, позволяющего анализировать динамику сцены и вычислять оптический поток — векторное поле, описывающее смещение пикселей между последовательными кадрами. Однако традиционные методы оценки оптического потока сталкиваются с трудностями при работе с большими смещениями, сложными текстурами и динамическими объектами.

В этой статье предлагается подход, основанный на применении метода LoFTR (Local Feature Transformer) для анализа оптического потока.

Обоснование выбора LOFTR

Классические методы оптического потока делятся на разреженные и плотные . Разреженные методы, такие как Lucas–Kanade [1], сначала выделяют характерные точки изображения (например, углы или участки с высокой текстурой), формируют их дескрипторы и устанавливают соответствия между кадрами. Плотные методы, такие как Horn–Schunck [2] и

Фарне́бэк [3], оценивают векторы движения сразу для всех пикселей изображения, без явного выделения ключевых точек.

Однако при слабой текстуре, повторяющихся паттернах или значительных изменениях освещения все эти подходы сталкиваются с проблемами: снижается точность оценки потока, а в случае сильных артефактов движение становится недостоверным. Это проявляется в виде ошибочных векторов, как показано на рисунке 1а и рисунке 1б, где зелёным цветом отображён найденный оптический поток

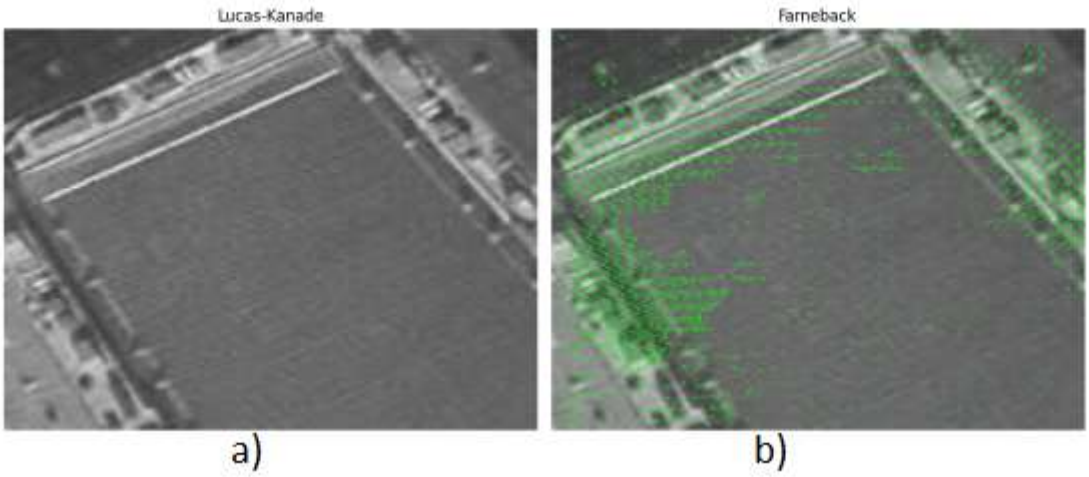


Рис. 1. Результаты работы алгоритмов а)Лукаса-Канаде б)Фарнебака

С развитием глубокого обучения альтернативным решением стало применение свёрточных нейронных сетей (CNN), способных извлекать признаки, устойчивые к шуму и изменениям условий съёмки. Однако даже эти методы сталкиваются с проблемами при обработке сложных сцен с большими смещениями или динамическими объектами, рисунок 2.

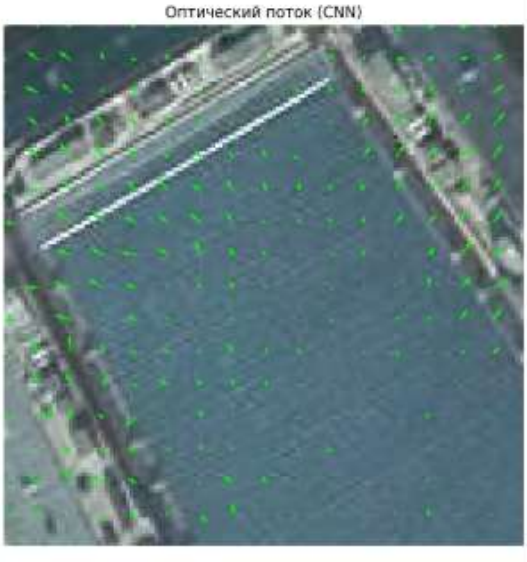


Рис. 2. Результаты работы CNN

На смену этим подходам пришёл LoFTR (Local Feature Transformer) — инновационная архитектура, отказавшаяся от традиционных детекторов признаков. LoFTR использует трансформеры с механизмами самовнимания и кросс-внимания [4] для обработки плотных локальных признаков, извлечённых из свёрточной сети. Сначала сопоставление проводится на

уровне признаков с низким разрешением (примерно 1/8 от исходного размера изображения), где выбираются наиболее достоверные соответствия. Затем эти точки уточняются до субпиксельной точности с помощью корреляционных методов.

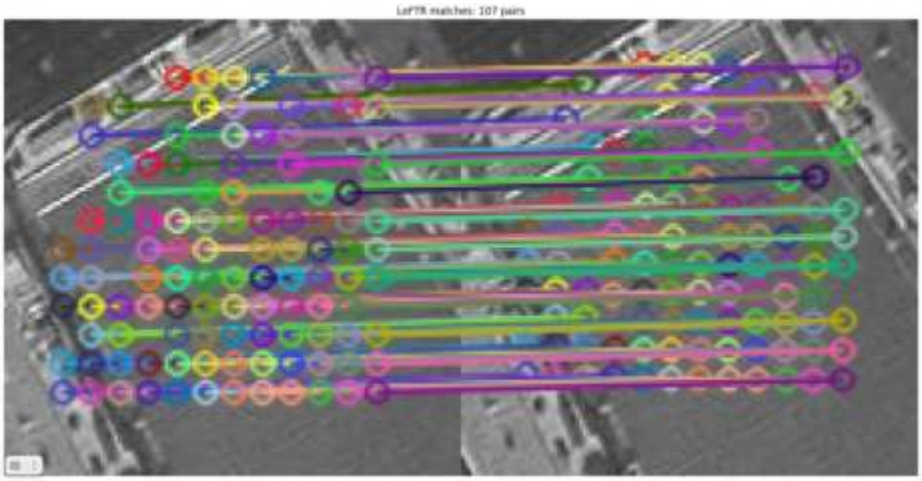


Рис. 3. Результаты работы LoFTR

Одним из современных подходов к поиску соответствий между кадрами является метод LoFTR [4], основанный на архитектуре трансформера. В отличие от классических методов, он обеспечивает высокую устойчивость к сложным условиям, таким как плохая текстура, повторяющиеся паттерны и изменения освещения.

Определения преобразования между изображениями

Для извлечения глобального движения сцены на основе локальных соответствий, найденных с помощью LoFTR, разработана нейронная сеть прямого распространения. Модель принимает на вход фиксированный набор из 100 точек, каждая из которых описывается четырьмя значениями: координатами в первом кадре и смещением между кадрами. Таким образом, размерность входного вектора составляет 400 элементов.

Архитектура сети представляет собой многослойный персептрон с двумя скрытыми слоями. На выходе модель предсказывает три параметра: угол поворота и компоненты сдвига по осям X и Y.

Обучение проводилось на синтетических данных с известными параметрами движения. В качестве функции потерь использовалась среднеквадратичная ошибка (MSE), измеряющая отклонение предсказаний от эталонных значений.

Практическая реализация и тестирование алгоритма

В рамках проведённого исследования было разработано программное обеспечение, предназначенное для восстановления траектории движения камеры на основе анализа видеопоследовательности. На вход система получает видеофрагмент, на выходе визуализируется предполагаемый маршрут, пройденный камерой.

Предложенный алгоритм показал стабильную работу в условиях простых сцен с плавным движением.

На рисунке 4б представлен эксперимент с прямолинейным движением: камера закреплена на смартфоне, перемещение на 4 м по ровной травянистой поверхности. Данные подтвердили точное восстановление векторов смещения при отсутствии динамических объектов.

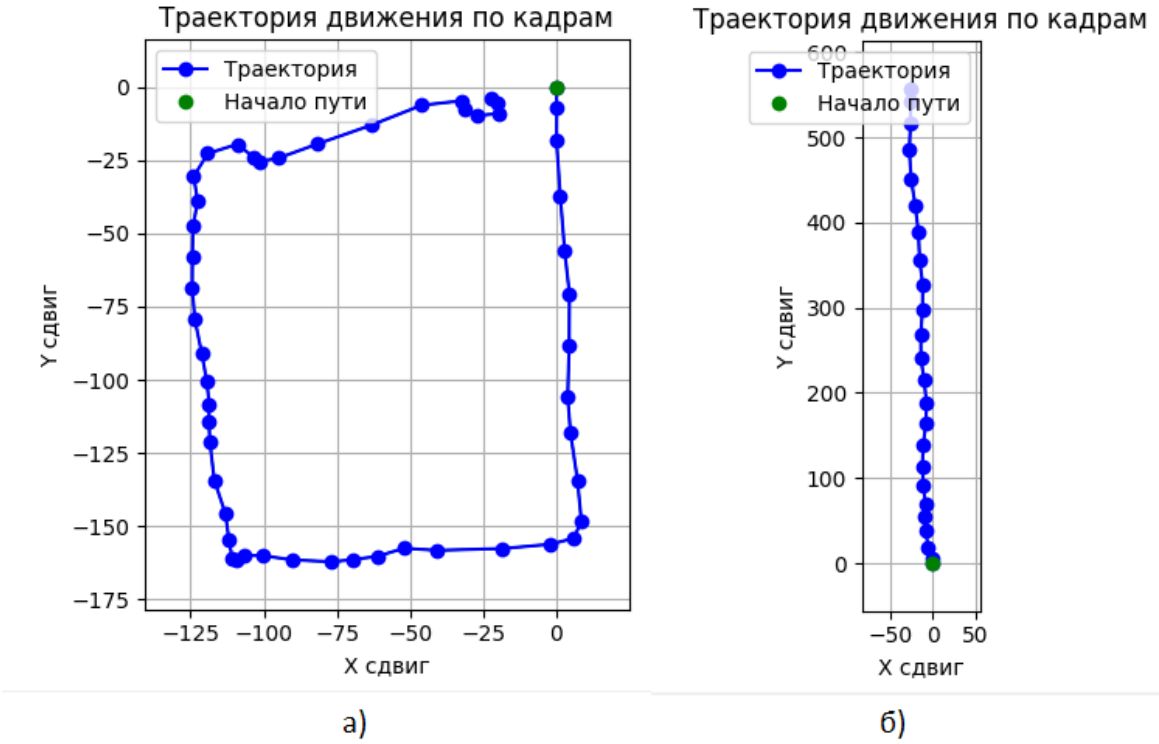


Рис. 4. Результаты работы алгоритмов а)прямое перемещение б)перемещение по квадрату

На рисунке 4а – тест на замкнутой траектории: система перемещалась по прямоугольнику и вернулась в исходную точку.

Вывод

Предложенный метод, сочетающий LOFTR для извлечения плотных соответствий и нейросетевую регрессию для оценки аффинных преобразований, демонстрирует высокую эффективность в задачах восстановления траектории движения камеры по видеопотоку без использования геолокации. Однако текущая реализация имеет ограничения при работе с динамическими сценами, включающими резкие повороты, вибрации устройства и скачкообразные перемещения, которые приводят к искажению траектории.

Список литературы

1. LUCAS B. D., KANADE T. An iterative image registration technique with an application to stereo vision // Proceedings of the 7th International Joint Conference on Artificial Intelligence. — 1981. — Vol. 2. — P. 674–679.
2. HORN B. K. P., SCHUNCK B. G. Determining optical flow // Artificial Intelligence. — 1981. — Vol. 17. — P. 185–203.
3. FARNEBÄCK G. Two-Frame Motion Estimation Based on Polynomial Expansion // Lecture Notes in Computer Science. — 2003. — Vol. 2749. — P. 363–370.
4. VASWANI A. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar et al. — Advances in Neural Information Processing Systems, 2017.
5. SUN J., LUO Z., MISHKIN D. et al. LoFTR: Detector-Free Local Feature Matching with Transformers // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2021. — P. 8922–8931. — DOI:10.1109/CVPR46437.2021.00885.
6. SUN D. PWC-Net: CNNs for Optical Flow Using Pyramid Warping, Cost Volume, and Regularization / D. Sun, X. Yang, M. Y. Liu et al. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2018. — arXiv:1709.02371.

7. ZHANG J. LoFTR: Detector-Free Local Feature Matching with Transformers / J. Zhang, Y. Tian, L. Wang et al. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2021. — arXiv:2104.00680.

КЛАССИФИКАЦИЯ МУЗЫКИ ПО ЭМОЦИЯМ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

М.К. Иванова

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Аннотация. В работе проведён анализ размеченного набора данных, включающего 900 аудиозаписей продолжительностью 30 с., каждая из которых относится к одному из четырёх эмоциональных классов: счастье, агрессию, подавленность и расслабленность. На первом этапе данные унифицируются и преобразуются в удобный формат, после чего из них извлекаются акустические признаки — традиционные MFCC и мел-спектрограммы, а также хромограмма, реже применяемый в подобных задачах музыкальный параметр. На втором этапе на основе сформированных признаков строятся и обучаются модели машинного обучения, далее оцениваются их классификационная точность и эффективность выбранных признаков и алгоритмов. Основная новизна исследования заключается в использовании хроматических признаков для распознавания эмоционального состояния по аудио, что демонстрирует их высокую информативность и потенциал для повышения качества классификации.

Ключевые слова: ML, MER (music emotion recognition), машинное обучение, классификация, метод, метрики, Recall, F1-Score, Precision, mel-спектрограмма, MFCC, хромограмма.

Подготовка данных и извлечение признаков

Актуальность данного исследования заключается в развитии технологий машинного обучения, персонализации музыкального контента и управлении им в интеллектуальной системе.

Эксперимент проводится на основе данных [1], на которые ссылаются авторы статьи “A Comparison Study of Deep Learning Methodologies for Music Emotion Recognition” [2], в которой описывалось решение той же задачи с помощью других методов машинного обучения. Данные являют собой пример размеченного набора аудиообразцов, каждый из которых отнесен лишь к одному из четырех кластеров. Эти кластеры представлены на рисунке ниже.

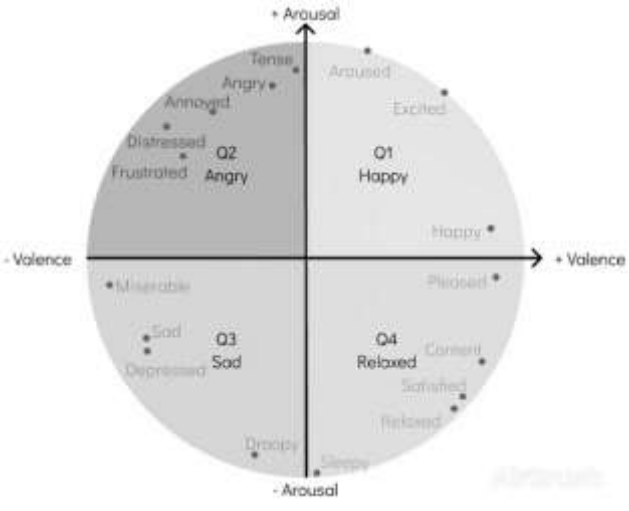


Рис.1 Разбиение данных на кластеры

При первичном анализе удается установить идеальную сбалансированность классов – благоприятный фактор при обучении модели.

Набор данных, на основе которого может быть осуществлено построение ML-моделей и их обучение, формируется после извлечения из каждого музыкального образца следующих значимых акустических признаков: Mel-спектрограммы – зависимости спектральной плотности мощности сигнала от времени, хромограммы – структуры, содержащей информацию об интенсивности 12 классов частот, MFCC (Мел-частотных кепстральных коэффициентов) – признака, характеризующего энергию спектра сигнала.

Построение моделей машинного обучения

Эксперимент включает в себя последовательное построение нескольких моделей машинного обучения и подбор гиперпараметров с помощью поиска по сетке GridSearchCV. Выбор моделей основан на их способности обрабатывать акустические признаки. Результаты обучения сведены в Таблицу 1.

Таблица 1

Результаты обучения ML-моделей

Модель	Квадрант	Precision (точность модели)
Logistic regression – логистическая регрессия, метод, основывающийся на предсказании вероятности принадлежности объекта к классу.	Q1	0.53
	Q2	0.75
	Q3	0.36
	Q4	0.34
	accuracy	0.49
KNN (K-Nearest Neighbors) – метод, производящий классификацию объектов в, основываясь на принципах “соседства”.	Q1	0.51
	Q2	0.73
	Q3	0.45
	Q4	0.50
	accuracy	0.54
SVC (Support Vector Classification) – метод, использующий гиперплоскости для разделения в многомерном пространстве классов. Эффективен при работе с большим количеством признаков.	Q1	0.52
	Q2	0.79
	Q3	0.40
	Q4	0.45
	accuracy	0.53
Decision Tree – подход на основе разбиения данных на более простые подмножества, эффективен при работе с числовыми и категориальными признаками.	Q1	0.43
	Q2	0.70
	Q3	0.34
	Q4	0.43
	accuracy	0.46
AdaBoost (Adaptive Boosting) – ансамблевый метод, в основе которого лежит последовательное обучение более слабых моделей и дальнейшее комбинирование результатов.	Q1	0.56
	Q2	0.81
	Q3	0.42
	Q4	0.50
	accuracy	0.56
	Q1	0.53

Модель	Квадрант	Precision (точность модели)
Gradient Boost – ансамблевый метод, основанный на последовательном построении моделей, каждая из которых сводит к минимуму ошибку предыдущей.	Q2	0.73
	Q3	0.39
	Q4	0.51
	accuracy	0.54
RFC (Random Forest Classifier) – метод, использующий множество деревьев решений для классификации. За счет обучения каждого дерева на случайной выборке данных переобучение сводится к минимуму.	Q1	0.61
	Q2	0.72
	Q3	0.45
	Q4	0.62
	accuracy	0.60

В процессе настройки моделей машинного обучения для исследования более широкого спектра гиперпараметров был применен метод halving в качестве замены традиционному поиску по сетке GridSearchCV. Данный метод представляет собой эффективный подход к оптимизации гиперпараметров и селекции моделей, который основывается на поэтапном распределении вычислительных ресурсов. Тем не менее, следует отметить, что использование указанного метода не ведет к улучшению метрик оценки моделей, что указывает на необходимость дальнейшего анализа и оптимизации процессов выбора гиперпараметров и моделей.

Анализируя полученные результаты, можно сделать вывод, что образцы, относящиеся к различным кластерам в ML-моделях, идентифицируются с разной степенью успеха. Метрики всех моделей лучше всего идентифицируют агрессию (второй квадрант). В модели адаптивного бустинга precision достигла 81%. Точнее остальных оказалась модель классификации, построенная на основе случайного леса, продемонстрировавшая при оптимизированных гиперпараметрах accuracy 60%. Неоднородность результатов для разных квадрантов и недостаточно высокие оценки, полученные в результате обучения всех моделей, могут свидетельствовать о неоптимально подобранном наборе признаков.

Для решения задачи MER дополнительную полезную информацию могут предоставить предварительно логарифмированные для выделения слабых нот хромограммы:

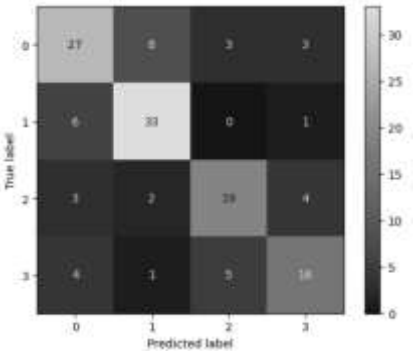
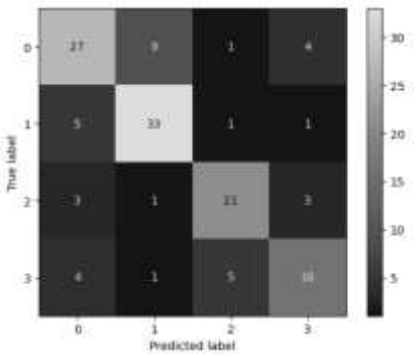
- CQT (Частотная Квадратурная Трансформация), охватывающая широкий диапазон частот,
- STFT (Кратковременное преобразование Фурье), учитывающее динамику атаки,
- CENS (Центрированный Энергетический Набор), устойчивый к шуму.

Добавление этих признаков, а также данных о темпе, в сочетании с ранее использованными MFCC и мел-спектрограммами формирует набор признаков для следующего эксперимента.

В качестве моделей для обучения на измененном наборе признаков выбираются две наиболее сильные: классификатор на основе случайного леса, лучше других справившийся с задачей при проведении предыдущего эксперимента, а также высокопроизводительный XGB-классификатор – ансамблевый метод машинного обучения, работающий путем создания случайного леса и обучения каждого дерева с учетом ошибок предыдущего. Для подбора параметров и настройки моделей используется инструмент для автоматической оптимизации гиперпараметров Optuna [3], алгоритмы которого базируются на следующих методах подбора: Grid Search (поиск по сетке), Random Search (случайный поиск), а также байесовская оптимизация. Метрики, полученные в результате обучения моделей, отражены в Таблице 2.

Таблица 2

Результаты обучения моделей на втором наборе признаков

Модель	Квадрант	precision	recall	f1-score	support	
RandomForestClassifier	Q1	0.68	0.61	0.67	41	
	Q2	0.75	0.82	0.79	40	
	Q3	0.70	0.68	0.69	28	
	Q4	0.67	0.62	0.64	26	
	accuracy	0.70				
XGBClassifier	Q1	0.69	0.66	0.68	41	
	Q2	0.75	0.82	0.79	40	
	Q3	0.75	0.75	0.75	28	
	Q4	0.67	0.62	0.64	26	
	accuracy	0.72				

Усиление исходного датасета за счёт добавления акустических хромограммных признаков, таких как CQT, STFT и CENS, в сочетании с использованием метода Optuna для подбора оптимальных гиперпараметров позволило существенно улучшить эффективность моделей, повысив метрики в среднем на 9-10%. При этом неоднородность предсказаний значительно уменьшилась, что может свидетельствовать о выборе для эксперимента признаков, более подходящих для работы с музыкальными аудиообразцами. По-прежнему построенные классификаторы лучше других идентифицирует агрессию (оценки для Q2 значительно выше, чем для остальных квадрантов). Наилучший результат продемонстрировала модель XGB-классификатора, которая показала высокую степень адаптивности к специфике рассматриваемых данных.

Список литературы

1. MOODetector Datasets and Extracted Features 4Q audio emotion dataset // MIR.cisuc URL: https://mir.dei.uc.pt/downloads.html (дата обращения: 14.04.2025).
2. A Comparison Study of Deep Learning Methodologies for Music Emotion Recognition // MDPI URL: https://www.mdpi.com/1424-8220/24/7/2201 (дата обращения: 17.03.2025).
3. Efficient Optimization Algorithms // Optuna URL: https://optuna.readthedocs.io/en/stable/tutorial/10_key_features/003_efficient_optimization_algorithms.html (дата обращения: 19.03.2025).

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ПРОГНОЗИРОВАНИЯ КАЧЕСТВА ВОЗДУХА НА ОСНОВЕ ГИБРИДНЫХ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ

ИМАНБАЕВ А.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Аннотация. В работе представлено исследование гибридной архитектуры нейронных сетей для прогнозирования концентрации загрязняющих веществ в городской среде. Предложенная модель объединяет сверточные нейронные сети (CNN) для анализа пространственных данных и сети с долгой краткосрочной памятью (LSTM) для обработки временных рядов. Экспериментальная оценка на данных 5 крупных городов России показала значительное улучшение точности прогнозирования (снижение RMSE на 37.4%) по сравнению с существующими методами. Результаты демонстрируют эффективность применения механизма пространственно-временного внимания для задач экологического мониторинга.

Ключевые слова: глубокое обучение, экологический мониторинг, сверточные нейронные сети, LSTM, качество воздуха

Введение

Проблема загрязнения воздуха в городах является одной из наиболее актуальных экологических проблем современности. Точное прогнозирование концентрации загрязняющих веществ представляет значительную научную и практическую ценность для минимизации рисков для здоровья населения [1]. Существующие подходы к прогнозированию имеют ограничения, связанные со сложностью учета множества факторов и неоднородной пространственно-временной динамикой процессов загрязнения [2].

Традиционные методы, такие как авторегрессионные модели (ARIMA) и алгоритмы машинного обучения (случайные леса, SVM), не учитывают в полной мере пространственные зависимости и нелинейные взаимосвязи между факторами [3]. Методы глубокого обучения потенциально способны преодолеть эти ограничения, однако требуют разработки специализированных архитектур.

В данной работе представлено исследование гибридной архитектуры нейронной сети для прогнозирования качества воздуха, объединяющей преимущества CNN и LSTM с интеграцией механизма пространственно-временного внимания.

Методология

Для исследования использованы данные с 85 станций мониторинга качества воздуха в 5 крупных городах России за 2020-2023 гг. Набор данных включает:

Почасовые измерения 6 загрязняющих веществ (PM2.5, PM10, NO_2, CO, SO_2, O_3)

Метеорологические параметры

Данные о дорожном трафике

Информация о промышленных выбросах

Общий объем выборки составил 5.7 миллионов записей. Для предварительной обработки применялись методы заполнения пропусков (k-ближайших соседей, скользящее среднее) и удаления выбросов.

Предложенная гибридная CNN-LSTM архитектура включает:

1. **Блок пространственного кодирования (CNN):** три сверточных слоя с фильтрами 3×3 для извлечения пространственных паттернов.

2. **Блок временного кодирования (LSTM):** два слоя LSTM для моделирования временных зависимостей.
3. **Механизм внимания:** пространственно-временной механизм внимания для определения значимости различных регионов и временных интервалов.
4. **Блок прогнозирования:** полносвязная сеть с остаточными соединениями для формирования итогового прогноза.

Результаты

Таблица 1 представляет результаты сравнения различных методов прогнозирования. Предложенная модель CNN-LSTM с механизмом внимания превосходит существующие подходы по точности на всех горизонтах прогнозирования.

Таблица 1

Сравнение точности прогнозирования (RMSE) для PM2.5

Метод	1 час	6 часов	12 часов	24 часа
ARIMA	8.92	14.87	22.35	31.68
Random Forest	7.81	13.24	19.47	28.31
MLP	7.53	12.65	18.92	27.45
CNN	6.78	11.52	17.84	25.69
LSTM	6.21	10.85	16.73	24.18
CNN-LSTM (без внимания)	5.46	9.38	14.87	21.93
CNN-LSTM с вниманием	**3.94**	**7.12**	**11.35**	**17.82**

Анализ результатов показывает, что:
1. Нейросетевые модели превосходят традиционные методы, особенно при увеличении горизонта прогнозирования.
2. Механизм внимания обеспечивает существенное улучшение точности (до 37.4% для 24-часового прогноза).

Анализ факторов влияния

Для определения относительной важности различных факторов проведен эксперимент с последовательным исключением групп признаков (таблица 2).

Таблица 2

Влияние факторов на точность прогнозирования (RMSE для PM2.5, 12 часов)

Конфигурация модели	RMSE	Изменение
Полная модель	11.35	-
Без метеоданных	15.87	+39.8%
Без данных о трафике	13.24	+16.7%
Без данных о промышленных выбросах	12.18	+7.3%
Без пространственного внимания	13.92	+22.6%
Без временного внимания	14.51	+27.8%

Метеорологические данные имеют наибольшее влияние на точность, что соответствует физике процессов распространения загрязнений. Значительное влияние также оказывают механизмы пространственного и временного внимания.

Пространственно-временной анализ ошибок

Наибольшие ошибки прогнозирования наблюдаются:

- В плотно застроенных районах со сложной геометрией
- Вблизи промышленных объектов с нерегулярными выбросами
- На границах между различными типами землепользования

Временной анализ выявил повышенную погрешность в периоды:

- Резких изменений метеорологических условий
- Утренних и вечерних пиков транспортного трафика
- Проведения строительных работ и действия других временных источников загрязнения

Заключение

В работе представлено исследование гибридной нейросетевой архитектуры CNN-LSTM с механизмом внимания для прогнозирования качества воздуха. Экспериментальная оценка показала значительное улучшение точности по сравнению с существующими методами.

Основные научные результаты:

1. Разработана специализированная архитектура, учитывающая пространственно-временную специфику данных о качестве воздуха.
2. Интегрирован механизм внимания для автоматического определения значимости различных факторов и регионов.
3. Проведена количественная оценка влияния различных групп факторов на точность прогнозирования.
4. Выявлены пространственно-временные паттерны в распределении ошибок.

Дальнейшие исследования будут направлены на разработку методов интерпретации решений нейросетевых моделей, интеграцию дополнительных источников данных и адаптацию модели для прогнозирования экстремальных событий загрязнения.

Список литературы

1. Wei J., Li Z., Cribb M. et al. Improved 1-km resolution PM2.5 estimates across China using enhanced space-time extremely randomized trees // Atmospheric Chemistry and Physics. – 2021. – Vol. 21. – P. 7863-7880.
2. Huang C.-J., Kuo P.-H. A deep CNN-LSTM model for particulate matter (PM2.5) forecasting in smart cities // Sensors. – 2018. – Vol. 18. – P. 2220.
3. Cheng W., Shen Y., Zhu Y. et al. A neural attention model for urban air quality inference: Learning the weights of monitoring stations // AAAI Conference on Artificial Intelligence. – 2022. – P. 5623-5631.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ОБУЧЕНИИ: ЛУЧШИЕ КЕЙСЫ ВНЕДРЕНИЯ В ШКОЛАХ И ВУЗАХ

КОРНОУХОВА Т.В., КОПЫЛОВА Д.А.

ФГБОУ ВО «Поволжский государственный университет телекоммуникаций и информатики»

Аннотация. Статья рассматривает опыт применения искусственного интеллекта (ИИ) в российских и зарубежных образовательных учреждениях. В условиях активной цифровой трансформации современной системы образования, ИИ открывает новые возможности для персонализированного обучения, автоматизации рутинных задач и анализа данных. В статье анализируются успешные примеры внедрения ИИ в школы и вузы, достигнутые результаты и перспективы развития.

Ключевые слова: искусственный интеллект, образование, цифровая трансформация, персонализированное обучение, автоматизация, анализ данных, виртуальные ассистенты, гибридный интеллект, образовательные технологии.

Введение

Современная система образования переживает период активной цифровой трансформации, где технологии ИИ играют ключевую роль. ИИ трансформирует образование, обеспечивая персонализацию учебного процесса, автоматизацию рутинных задач и аналитику данных для повышения качества обучения. По данным ЮНЕСКО, технологии ИИ способствуют инклюзивности и доступности образования, но требуют строгого этического регулирования для защиты данных и сохранения человеческого фактора [1].

Московская электронная школа: адаптивное обучение

Проект «Московская электронная школа» (МЭШ) — один из ведущих примеров применения ИИ в российских школах. Платформа использует алгоритмы машинного обучения для анализа успеваемости учеников и подбора индивидуальных заданий, что повышает вовлеченность. В 2023 году МЭШ охватывала около 1 миллиона учеников в примерно 800 школах Москвы — около 50% от общего числа школ города — и предоставляла доступ к более чем 1,5 миллионам цифровых ресурсов, включая интерактивные уроки и тесты [2]. Педагоги отмечают сокращение времени на подготовку уроков благодаря автоматизации, а аналитические инструменты помогают выявлять образовательные потребности учеников. Успех МЭШ обусловлен интеграцией ИИ с облачными технологиями и постоянным обновлением на основе обратной связи.

Университет ИТМО: персонализированные траектории

Университет ИТМО в Санкт-Петербурге применяет ИИ для создания индивидуальных образовательных маршрутов, анализируя академические результаты, навыки и интересы студентов, чтобы рекомендовать курсы и проекты. Университет сотрудничает с международными платформами, такими как Coursera и edX, предоставляя доступ к множеству образовательных модулей [3]. ИИ помогает выявлять академические трудности на ранних этапах, снижая количество пересдач и повышая удовлетворенность студентов. Подход ИТМО укрепляет его позиции в цифровом образовании за счет сочетания локальных инноваций и глобальных ресурсов.

Рособрнадзор: обеспечение объективности ЕГЭ

Федеральная служба по надзору в сфере образования и науки (Рособрнадзор) использует ИИ для мониторинга Единого государственного экзамена (ЕГЭ). Нейросети анализируют видеопотоки с экзаменационных пунктов, выявляя нарушения, такие как использование шпаргалок, охватывая большинство из 85 регионов России в 2023 году [4]. ИИ также автоматизирует проверку письменных работ, повышая эффективность и согласованность оценок. Эти меры оптимизируют ресурсы и укрепляют доверие к объективности ЕГЭ.

Технологический институт Джорджии: виртуальный ассистент Jill Watson

С 2016 года в Технологическом институте Джорджии используется ИИ-ассистент Джилл Уотсон, созданный на основе IBM Watson. Он помогает студентам онлайн-курса "Методы искусственного интеллекта", отвечая на вопросы и проверяя задания. Профессор Эшок Гёль внедрил эту систему, чтобы справиться с большим потоком запросов — около

10 000 вопросов за семестр. ИИ обучался на данных учебного форума и со временем достиг 97% точности ответов [5]. Интересно, что студенты не сразу поняли, что общаются с программой. Джилл Уотсон успешно справлялась с типовыми вопросами, освобождая преподавателей от рутины [6].

Университет Стэнфорда

Исследователи из Стэнфорда разработали программу на основе машинного обучения, которая помогает студентам, сталкивающимся с трудностями в самостоятельном цифровом обучении. Программа не только предсказывает, когда студент может начать испытывать трудности, но и рекомендует решения, такие как возврат к предыдущим урокам или советы по механике игры. Программа была протестирована в сотрудничестве с War Child Holland, организацией, которая предоставляет планшеты и обучающее ПО школам в зонах конфликтов. Исследователи использовали данные о достижениях 1,170 угандийских школьников, чтобы обучить модель. Модель смогла предсказать, когда ученик начнет испытывать трудности, и в четырех из шести случаев рекомендации системы совпадали с рекомендациями эксперта-человека.

Цель программы — не заменить учителей, а помочь им, предоставляя рекомендации, когда студент сталкивается с трудностями. Это может быть полезно как в бедных, так и в богатых странах, а также для улучшения профессионального обучения и повышения квалификации взрослых [7].

Проблемы и перспективы развития

Несмотря на значительные успехи, внедрение ИИ в образование сопряжено с рядом сложностей. К технологическим вызовам относятся необходимость регулярного обновления алгоритмов и проблемы интеграции с устаревшими системами. Педагогические сложности включают сопротивление части преподавателей и риск чрезмерной автоматизации учебного процесса. Особого внимания заслуживают этические вопросы, связанные с конфиденциальностью данных и проблемой алгоритмической предвзятости.

Для успешного внедрения ИИ в образование необходимо учитывать следующие аспекты. Во-первых, обеспечить технологическую совместимость новых ИИ-систем с существующими образовательными платформами и базами данных. Во-вторых, проводить регулярные тренинги и семинары для преподавателей по использованию ИИ-инструментов в учебном процессе. В-третьих, разработать и внедрить этические стандарты использования ИИ в образовании, включая защиту данных и предотвращение алгоритмической предвзятости. В-четвертых, сохранять баланс между автоматизацией и индивидуальным подходом к обучению, чтобы избежать чрезмерной стандартизации образовательного процесса.

Перспективы развития ИИ в сфере образования включают создание эмоционального искусственного интеллекта для анализа вовлечённости учащихся, разработку комплексных цифровых помощников преподавателя, совершенствование систем прогнозной аналитики в реальном времени, а также интеграцию с технологиями виртуальной и дополненной реальности. Особенно перспективной представляется модель «гибридного интеллекта», которая сочетает преимущества искусственного интеллекта с опытом преподавателей [8].

Заключение

Проведенный анализ демонстрирует, что российские образовательные учреждения успешно внедряют технологии искусственного интеллекта, достигая результатов,

сопоставимых с ведущими зарубежными аналогами. Опыт как российских, так и зарубежных учебных заведений подтверждает, что ИИ становится важным инструментом повышения эффективности образовательного процесса. Успешная реализация требует комплексного подхода, учитывающего технологические, педагогические и организационные аспекты. Дальнейшее развитие образовательных технологий должно быть направлено на создание сбалансированных систем, оптимально сочетающих возможности искусственного интеллекта с профессиональными компетенциями преподавателей.

Список литературы

1. Guidance for generative AI in education and research [Электронный ресурс]. – Режим доступа: https://unesdoc.unesco.org/ark:/48223/pf0000386693
2. Искусственный интеллект поможет школьникам изучить вероятность и статистику [Электронный ресурс]. – Режим доступа: https://edupressa.vm.ru/gazeta/plody-prosvescheniya/iskusstvennyj-intellekt-pomozhet-shko/
3. Mass & Quality: How Russia and China Transform Education With AI [Электронный ресурс]. – Режим доступа: https://news.itmo.ru/en/education/cooperation/news/13589/
4. ИИ обучили следить за нарушениями во время сдачи ЕГЭ [Электронный ресурс]. – Режим доступа: https://habr.com/ru/news/783600/
5. Meet Jill Watson: Georgia Tech's first AI teaching assistant [Электронный ресурс]. – Режим доступа: https://pe.gatech.edu/blog/meet-jill-watson-georgia-techs-first-ai-teaching-assistant
6. Студенты не смогли отличить консультацию искусственного интеллекта от ассистента профессора [Электронный ресурс]. – Режим доступа: https://life.ru/p/407510
7. Stanford researchers created a program to help when students get stuck in self-paced digital learning [Электронный ресурс]. – Режим доступа: https://hai.stanford.edu/news/using-artificial-intelligence-understand-why-students-are-struggling
8. Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development [Электронный ресурс]. – Режим доступа: https://roscongress.org/en/materials/iskusstvennyy-intellekt-v-obrazovanii-problemy-i-vozmozhnosti-dlya-ustoychivogo-razvitiya/

МЯГКИЕ ИЗМЕРЕНИЯ НЕЙРОНЕЧЁТКИХ СОСТАВЛЯЮЩИХ XAI 2.0: МУЛЬТИУРОВНЕВАЯ ГИБРИДНАЯ АРХИТЕКТУРА (CNN + ANFIS)

Ю.В Трофимов, Е.М. Кузнецов, И.П. Муравьев

Государственный университет «Дубна»

Аннотация. В статье рассматривается многоуровневый подход к объяснительному искусственному интеллекту для диагностики пневмонии в рентгенографии грудной клетки. Использование глубоких нейронных сетей и компьютерного зрения в медицине требует высокой точности и объяснимости результатов для специалистов. В рамках исследования интегрированы методы объяснительного ИИ: для локальной интерпретации - Grad-CAM, LIME, для глобальной интерпретации - нейро-нечёткая система ANFIS, что позволяет визуализировать важные области на снимках и формировать продукционные правила. Также разработан модуль генерации текстовых заключений. Разработка ориентирована на интеграцию в клинические системы, с возможностью масштабирования и гибкой настройки под различные сценарии применения. Работа направлена на повышение доверия к алгоритмам ИИ в медицине через улучшение их прозрачности.

Ключевые слова: XAI 2.0, объяснимый искусственный интеллект, CNN, ANFIS, Grad-CAM++, Rule Aggregation, LIME, BloodMNIST, нечеткая логика, функциональная безопасность, интерпретируемость, медицинская диагностика, кластеризация K-Means, Focal Loss, нейронечёткие системы, мультиуровневая интерпретация.

Введение

В современных исследованиях компьютерного зрения и искусственного интеллекта (ИИ) особую значимость приобретает совокупность технологий, ориентированных на объяснимость и интерпретируемость (XAI — eXplainable AI). Классические глубокие свёрточные сети (CNN) позволяют достичь выдающихся результатов при решении задач, связанных с высокоразмерными изображениями, но традиционно остаются «чёрным ящиком», осложняющим процессы сертификации в контексте функциональной безопасности и информационной надёжности. На фоне ужесточающихся требований к прозрачности и воспроизводимости (особенно в клинической диагностике, промышленной автоматизации и смежных отраслях) возникает запрос на механизмы мягкой (soft) интерпретации принятых решений. В данном аспекте нейронечёткие системы (ANFIS) интегрируются с глубинными архитектурами, обогащая их формальным аппаратом нечеткой логики, а именно — функциям принадлежности (MF), субтрактивным методам кластеризации центров (K-Means), а также логическими правилами вида «Если-то» (fuzzy rule-based inference).

При этом достигается мультиуровневый подход к объяснимости (XAI 2.0), где несколько граней интерпретации — от локальной (LIME-подобный анализ активности правил) до глобальной (Grad-CAM++ для CNN, Rule Aggregation и визуализация решающей поверхности в нейронечётком пространстве) — усиливают общую транспарентность системы [1]. Использование нескольких уровней объяснимости призвано повысить робастность общей системы, удовлетворяя тем самым ключевым критериям доверия, связанным с верификацией и валидацией. Данная особенность формирует прочную основу для интеграции данных алгоритмов в контуры функциональной безопасности и информационной надёжности, где объяснимость является не менее важным аспектом, чем точность.

Одним из центральных кейсов применения предложенного гибридного решения является классификация клеток крови по медицинским изображениям. В качестве тестового стенда использован датасет BloodMNIST из семейства MedMNIST [2-4], содержащий высококачественные анотированные снимки различных типов клеток крови. Этот набор данных является частью медицинского стандартизированного тестового окружения и широко применяется для апробации моделей глубокого обучения в области гематологической диагностики. Поставленная задача включает распознавание нескольких классов клеток крови, что имеет критическое значение в медицинском анализе, так как позволяет выявлять патологические аномалии и автоматизировать предварительный этап диагностики. В этом контексте объяснимость модели играет первостепенную роль, поскольку клинические специалисты должны понимать, на основе каких признаков и закономерностей был вынесен тот или иной диагноз

Основная часть

Предлагаемая архитектура организована в виде двухэтапного конвейера. Первый блок — глубокая сверточная нейронная сеть, состоящая из последовательности свёрточных модулей с BatchNorm и Dropout, объединённых пуллинговыми слоями, что обеспечивает качественную редукцию пространственных размерностей и выделение дистинктивных признаков (feature maps). Для улучшения сходимости используется OneCycleLR, а механизмы вроде Gradual Unfreeze позволяют поэтапно «размораживать» верхние слои и тонко настраивать их под целевую задачу, сохраняя при этом общие весовые распределения

на низких слоях [5]. Проблема дисбаланса классов устраняется с помощью Focal Loss (в сочетании с параметром гамма) и Weighted Random Sampler, что гарантирует надлежащий учёт труднообучаемых и редко встречающихся примеров.

Второй блок — ансамбль нейронечётких систем (One-vs-All ANFIS), каждая из которых ориентирована на бинарное решение «принадлежит ли объект определённому классу». Вектор высокоуровневых признаков (извлечённых CNN) подаётся на вход ANFIS, где инициализация функций принадлежности (MF) выполняется с помощью K-Means кластеризации по каждому входному признаку. Фактически это реализация мягких измерений (fuzzy measurements) признаков, когда для каждой компоненты выделяется несколько гауссовых (или иных) функций принадлежности с автоматически определяемыми центрами и сигмами. Далее формируются правила в декартовом произведении MF, и линеаризованный вывод (Sugeno-типа) оптимизируется через мини-батч AdamW (для параметров MF) и метод наименьших квадратов (LSE) для линейных коэффициентов. Такая конфигурация обеспечивает гибридную интерпретируемость: в то время как CNN решает высокоуровневую задачу выделения признаков, нейронечёткая логика ANFIS декодирует их с позиции логических конструкций «если-то», которые верифицируются и могут быть агрегированы (Rule Aggregation) для устранения дублирующих либо коррелированных правил.

Значимость мультиуровневого объяснения в XAI 2.0 становится особенно очевидной при рассмотрении различных механик интерпретации в данном гибридном решении. На уровне CNN используется Grad-CAM++, визуализирующий регионы активации на карте признаков, отвечающие за наиболее весомый вклад в итоговую классификацию. Этот слой интерпретации даёт информацию об «области внимания» сети и крайне востребован в задачах, где необходимо проверить, что алгоритм основывается на релевантных фрагментах изображения (например, при медицинской диагностике).

На уровне ANFIS задействуются Rule Aggregation, подсветка весов firing strengths (сырых и нормированных) для каждого правила и LIME-подобные процедуры, которые позволяют наглядно объяснять, почему именно сработало конкретное сочетание функций принадлежности. Базирующаяся на нескольких MF для каждого признака матрица правил формирует нечеткую интерпретацию, характерную для прикладной области (допустим, «признак 1: высокий», «признак 2: слегка отрицательный», и т.д.). Далее пользователь или эксперт может проследить, как конкретные значения входных характеристик повлияли на уровень истинности каждого правила, какой совокупный вклад они внесли в решающее уравнение ANFIS и к какому итогу привели при объединении в One-vs-All схеме.

Комплексная мультиуровневая интерпретация — от визуализации CNN-карт активации (Grad-CAM++) до локальной (LIME-like) и глобальной (объединённая Rule Aggregation) нечёткой логики — способствует повышению прозрачности и доверия к решению. Это особенно актуально в задачах, где требуется соответствие не только высоким метрикам точности, но и строгим стандартам доказуемой функциональной безопасности (например, DIN EN ISO 13849 или IEC 61508) и информационной надёжности. Совмещение подобных механизмов XAI усиливает робастность системы, так как каждый уровень объяснимости дополняет друг друга, раскрывая аспекты принятия решений как на низком (уровень признаков), так и на высоком (уровень логических правил) уровне.

Выводы

В ходе проведённой исследования показано, что гибридная архитектура CNN + ANFIS, обогащённая мультиуровневыми механизмами объяснимости, способна объединять высокую точность классификации с расширенной интерпретируемостью. Удачная комбинация субтрактивной кластеризации (KMeans) при инициализации Membership Functions, мини-батч оптимизации в ANFIS (через AdamW и LSE) и Gradual Unfreeze в CNN создаёт оптимальный баланс между выразительностью представлений и контролем переобучения. Использование Focal Loss и Weighted Random Sampler минимизирует риски негативного влияния дисбаланса классов, а метрики ECE и ROC-AUC обеспечивают оценки надёжности и калиброванности итоговых вероятностных предсказаний.

Главной особенностью разработанной системы можно считать слияние нескольких уровней XAI: Grad-CAM++ даёт наглядное понимание «регионов внимания» сети, тогда как Rule Aggregation и LIME-подобная локальная интерпретация в ANFIS предоставляют пользователю дополнительные «точки входа» для анализа того, почему система сделала именно такой вывод и какие нечёткие правила в этом задействованы. Использование нескольких уровней объяснимости призвано повысить общую робастность и доверие к итоговому решению, что вписывается в растущие требования функциональной безопасности и информационной надёжности во всех сферах применения (медицинской диагностике, промышленной автоматизации, транспортной логистике и т.д.). Совокупность полученных результатов формирует предпосылки для более масштабного внедрения гибридных нейронечётких технологий, наглядно демонстрируя их потенциал для высокоточных и интерпретируемых систем класса XAI 2.0.

Список литературы

1. Аверкин А. Н. Объяснимый искусственный интеллект как часть искусственного интеллекта третьего поколения // Речевые технологии/Speech Technologies. 2023. №1. URL: https://cyberleninka.ru/article/n/obyasnimyy-iskusstvennyy-intellekt-kak-chast-iskusstvennogo-intellekta-tretiego-pokoleniya (дата обращения: 11.02.2025).
2. Jiancheng Yang, Rui Shi, Donglai Wei, Zequan Liu, Lin Zhao, Bilian Ke, Hanspeter Pfister, Bingbing Ni. Yang, Jiancheng, et al. "MedMNIST v2-A large-scale lightweight benchmark for 2D and 3D biomedical image classification." Scientific Data, 2023.
3. Jiancheng Yang, Rui Shi, Bingbing Ni. "MedMNIST Classification Decathlon: A Lightweight AutoML Benchmark for Medical Image Analysis". IEEE 18th International Symposium on Biomedical Imaging (ISBI), 2021.
4. Jiancheng Yang, Rui Shi, Donglai Wei, Zequan Liu, Lin Zhao, Bilian Ke, Hanspeter Pfister, & Bingbing Ni. (2024). [MedMNIST+] 18x Standardized Datasets for 2D and 3D Biomedical Image Classification with Multiple Size Options: 28 (MNIST-Like), 64, 128, and 224 (3.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10519652
5. Аверкин Алексей Николаевич Гибридные модулярные нейронные сети // Статистика и экономика. 2016. №4. URL: https://cyberleninka.ru/article/n/gibridnye-modulyarnye-neyronnye-seti (дата обращения: 05.01.2025).

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № 124112200072-2).

ОЦЕНКА ПРОИЗВОДИТЕЛЬНОСТИ И ЭФФЕКТИВНОСТИ ДИСТИЛИРУЕМЫХ, КВАНТУЕМЫХ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ С ОТКРЫТЫМ ИСХОДНЫМ КОДОМ ДЛЯ РАЗРАБОТКИ АГЕНТОВ ИИ НА ВСТРАИВАЕМЫХ УСТРОЙСТВАХ

Левушкина С.И, Панфилова Е.Д.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)»

Аннотация. В данной работе проводится сравнительный анализ больших языковых моделей с открытым исходным кодом с точки зрения их применимости для разработки интеллектуальных агентов. LLM были развернуты в Docker-контейнере, что позволило ограничить ресурсы, и протестированы. После прохождения тестов были выявлены лучшие модели для развертывания на встраиваемых устройствах.

Ключевые слова: большие языковые модели (LLM), интеллектуальные агенты, встраиваемые устройства, граничные вычисления (Edge Computing), Docker, DeepSeek-coder, YandexGPT-5-Lite, Mistral, TinyLlama.

Введение

В последние годы происходит быстрое развитие ИИ ассистентов во всех сферах жизни. Большинство существующих решений основано на больших языковых моделях, которые требуют значительные вычислительные ресурсы для работы. Перенос возможностей больших языковых моделей в среду ограниченных вычислений решает сразу несколько проблем:

1. Минимизация задержки при выдаче ответов при локальных вычислениях
2. Возможность работы оффлайн
3. Большая безопасность данных, так как нет необходимости их передачи при локальных вычислениях

Поиск моделей с лучшим соотношением качества генерации ко времени генерации на ограниченных ресурсах важная задача, которая поможет наладить повсеместное внедрение ИИ-ассистентов даже в ограниченных вычислительных средах, на встраиваемых устройствах [1].

Для решения такой задачи было выбрано несколько LLM, которые были развернуты на ограниченных устройствах и протестированы.

Развертывание моделей

Для тестирования моделей в условиях ограниченных ресурсов был создан докер-контейнер с возможностью дробления на уровне дравейров доступа к GPU и интегрированным фреймворком ollama [2]. Использовалась видеокарта Nvidia A5000 на 48 Gb памяти, ОС Ubuntu 20.04. Контейнер запускает Ollama - платформу, которая позволяет локально запускать большие языковые модели. Выделено 8 Гб видеопамяти GPU. Данные (модели, кэш, конфигурация) сохраняется в папку на хосте. В контейнере предустановлены драйвера nvidia [3], позволяющие отслеживать и ограничивать потребление видеопамяти.

В этом контейнере были развернуты и запущены три модели, чтобы измерения точности были более достоверными за счет одинаковых вычислительных ресурсов. Запуск производился с дополнительным флагом --verbose для вывода отладочной информации от Ollama.

Тестирование

Для тестирования моделей был написан универсальный (для любой модели, развернутой в контейнере) bash-скрипт, который в качестве дополнительного параметра при вызове принимает название тестируемой модели.

В качестве бенчмарков были выбраны несколько промптов:

“Напиши на C++ алгоритм сортировки методом пузырька” - пример, который показывает, как модель пишет код. Правильность написания кода оценивалась автоматически, сухость и понятность ответа вручную.

“Объясни Docker простыми словами” - показывает, насколько емко и информативно модель может дать информации по запросу без наводящих вопросов.

“Реши уравнение: $7x^3 + 3x^2 - x - 9 = 0$” - показывает умение модели решить простейшую математическую задачу, а также подробность и понятность выдаваемого решения.

“Напиши текст длиной ровно 500 токенов о преимуществах ИИ” - показывает умение модели “рассуждать” на какую-либо тему.

Первый-третий бенчмарки запускались по три раза для каждой модели, для итоговой оценки из трех результатов бралось среднее значение времени, скорости генерации и точности.

Четвертый бенчмарк запускался по одному разу для каждой модели, так как оценивался человеком.

Описание моделей

Первой рассмотренной моделью стала Mistral, которая включает в себя 7 миллиардов параметров, она мощная, но компактная. Такая модель хорошо подходит для развертывания на ограниченных ресурсах. Модель поддерживает английский и базовый русский языки. В модели оптимизировано использование памяти, что позволяет ей работать на 8 Гб оперативной памяти. Использование VRAM до 8 Гб.

DeepSeek-coder — это специализированная модель для написания кода, которая имеет 6.7 миллиардов параметров. Модель поддерживает множество языков программирования, хорошо работает на английском языке, формально есть поддержка русского языка. Ответы на русском менее качественны, чем в других моделях.

Модель требует 6 Гб оперативной памяти, поэтому был использован докер-контейнер с 8Гб RAM, как и при развертывании предыдущих моделей. Требуется 7 Гб видеопамяти.

TinyLlama 1.1B - компактная языковая модель с 1.1 миллиардом параметров. Требует около 1.5-2 ГБ оперативной памяти. Поддерживает английский и русский языки, но качество русскоязычных ответов крайне низкое. Модель распространяется под лицензией Apache 2.0. Разработана на базе архитектуры Llama.

Модель YandexGPT 5 Lite имеет 8 миллиардов параметров и идеально подходит для работы на русском языке. Также у модели есть поддержка английского языка. Для корректной работы модели требуется 30 Гб оперативной памяти. Для оценки этой модели развернем ее в докер-контейнере с достаточными ресурсами и проведем те же измерения, что и для других.

Модель показала лучшие скорость генерации и точность ответов. Это объяснимо тем, что, во-первых, модель была развернута в контейнере с большими ресурсами, а во-вторых - тем, что тестирование проходило на русском языке, а YandexGPT 5 Lite имеет лучшую поддержку русского по сравнению с остальными испытуемыми моделями.

Таблица 1

Сравнение показателей моделей

Характеристика	**Mistral**	**YandexGPT 5 Lite**	**DeepSeek-coder**	**TinyLlama**

Количество параметров	7 млрд	8 млрд	7 млрд	1.1 млрд
Поддержка языков	Английский, базовый русский	Русский, английский	Английский, слабый русский	Английский, слабый русский
Требуемая RAM (Гб)	8	30	6	1.5
Требуемая VRAM (Гб)	8	30.1	7	1.6
Скорость генерации (токен/сек)	40	45	30	100
Задержка первого токена	1.2 с	1.8 с	1 с	<500 мс
Точность (бенчмарки)	84%	94%	73%	55%
Оптимизация под русский язык	Базовая	Лучшая	Слабая	Крайне слабая

Как и ожидалось, в тестах на русском языке лучше всего себя показала модель YandexGPT, так как она была развернута в среде с большим количеством ресурсов в отличие от остальных моделей. Потребление памяти максимальное среди рассмотренных моделей. Текст осмысленный, понятный, при ответах на вопросы, не требующие пояснений, модель дает лаконичный ответ без предисловий.

Модель deepseek-coder имеет средние время ответа и точность. При проверке ответов было замечено, что модель иногда теряет связь предложений при рассуждениях, что неудивительно, ведь модель предназначена для решения технических задач. Но в модель может легко давать технические определения, вычисления. Также модель лучше остальных справилась с написанием кода, качественно выполнив техническую часть задания.

Mistral показала себя лучше, чем deepseek: скорость генерации выше, точность больше. В ответах модель немногословна, даже при запросе генерации длинных рассуждений, ответ получается коротким. Иногда модель отвечает на английском языке. Но при этом точность все равно является самой высокой при меньших ресурсах.

Модель TinyLlama оказалась слабее всех - при наивысшей скорости генерации и наименьшем потреблении памяти тесты были провалены - модель не способна решить простейшую математическую задачу и дать осмысленный ответ на какой-либо промпт на русском языке, ответы модели на русском языке плохо читаемы (грамматика, формы слов спутаны).

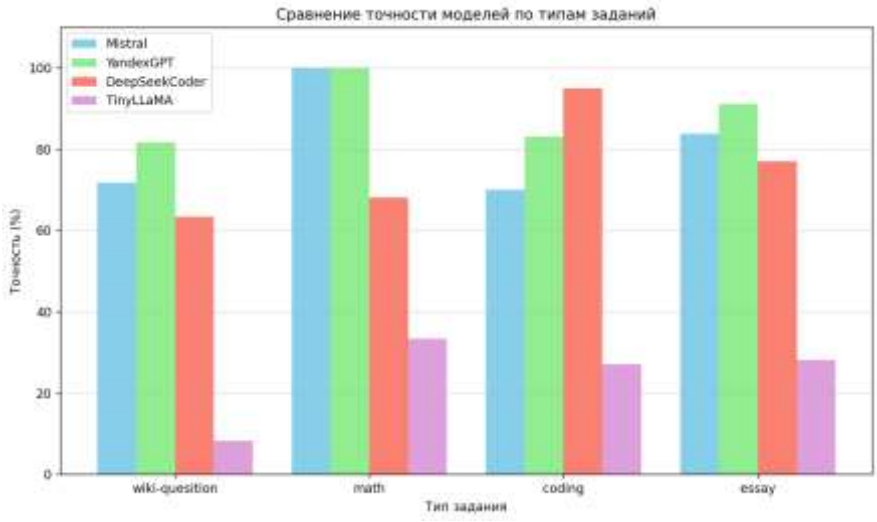


Рис. 1. Сравнение точности моделей по типам заданий

Выводы

По результатам исследования ясно, что развертывание больших языковых моделей в условиях ограничения ресурсов на ПК возможно, но либо с потерей качества генерации, либо с потерей скорости и необходимостью выделения больших ресурсов, из чего можно выдвинуть предположение, что на встраиваемых устройствах модели будут вести себя примерно так же.

В последние годы рынок ИИ-ассистентов быстро развивается и модели, которые еще недавно лидировали в производительности и качестве ответов уже на порядок уступают новым, появляющимся всё более часто и развивающимися все большими темпами. Поэтому внедрение в конечные устройства ИИ-агентов на основе устаревших нейросетей не имеет смысла и может отпугнуть будущих пользователей.

Решением проблемы может быть во-первых - использование моделей средней размерности (2-5 млрд параметров) и применение к ним техники донастройки и дообучения для определенных задач, что поможет сэкономить вычислительные ресурсы, получив при этом достаточное улучшение качества ответов моделей для из дальнейшего комфортного использования.

Так как исследования проводились на ПК с ограничением ресурсов, в дальнейшем имеет смысл провести подобные эксперименты со сравнением моделей уже на реальных устройствах [4], например таких, как Rocketchip (архитектура RISC-V) и более детально оценить качество моделей, используя популярные бенчмарк-сервисы, что потребует больше времени и ресурсов.

Список литературы

1. Daniel Situnayake, Jenny Plunkett. AI at the Edge: Solving Real-World Problems with Embedded Machine Learning
2. Hostkey [Электронный ресурс]. - URL: https://hostkey.ru/documentation/technical/gpu/ollama
3. Github [Электронный ресурс]. - URL: https://github.com/clearml/clearml-fractional-gpu
4. Gian Marco Iodice. Deep Learning for Embedded Systems

АНАЛИЗ СРЕД ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ МНОГОАГЕНТНЫХ СИСТЕМ РЕАЛЬНОГО ВРЕМЕНИ

Литягин С.М.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Аннотация. Проведен анализ сред имитационного моделирования многоагентных систем реального времени (СИМ МАС РВ), позволяющих отрабатывать интеллект агентов в сценарии командного взаимодействия и противодействия в динамическом, непрерывно меняющемся мире. Определены ключевые требования к таким средам, включая поддержку различных типов агентов, масштабируемость, моделирование в реальном времени и межагентной коммуникации. Выполнено сравнительное исследование четырех платформ: Gazebo, MAgent, RoboCup Soccer Simulator и Unity ML-Agents, выявившее их ограничения в контексте поставленных задач. Показано, что существующие решения не полностью удовлетворяют требованиям СИМ МАС РВ, что обосновывает необходимость разработки специализированной модульной среды.

Ключевые слова: интеллектуальный агент, многоагентная система, среда имитационного моделирования, реальное время

Введение

Современные исследования в области искусственного интеллекта всё чаще ориентированы на моделирование поведения интеллектуальных агентов, способных действовать в команде, координировать свои действия и адаптироваться к динамически изменяющимся условиям открытой среды реального времени. Такие интеллектуальные агенты в настоящее время используются в системах, функционирующих как в реальном мире (автономные роботы, беспилотные летательные аппараты и т.п.), так и в виртуальных средах (компьютерные игры и др.). Однако обучение и тестирование таких систем в реальном мире затруднено из-за высокой стоимости оборудования, рисков повреждения техники и ограничений в масштабируемости экспериментов.

Для преодоления этих ограничений широко применяются среды имитационного моделирования многоагентных систем реального времени. Тем не менее, большинство существующих решений либо ориентированы на физически достоверное моделирование, включая графику и динамику, либо заточены под ограниченные сценарии. При этом существующие среды, позволяющие разрабатывать и тестировать координированное поведение интеллектуальных агентов в командах с внутренней структурой и коммуникацией, имеют ограничения. Такие среди особенно важны на этапах разработки интеллектуального агента, когда интерес представляет интеллектуальное поведение агента, а не высокая точность моделирования динамики и графики, моделирование низкоуровневых сенсоров (лидарные облака, RGB-изображения с камеры и т.п.).

В работе анализируются среды имитационного моделирования многоагентных систем в реальном времени, предназначенные для обучения и тестирования интеллектуальных агентов, действующих в составе команд в условиях противодействия.

Объектом исследования является процесс моделирования командного взаимодействия агентов в условиях противодействия. Предметом исследования выступает среда имитационного моделирования многоагентных систем реального времени, обеспечивающая моделирование поведения агентов в рамках заданного сценария, сенсорного восприятия агентов и межагентной коммуникации.

Цель исследования заключается в выделении преимуществ и недостатков существующих сред имитационного моделирования многоагентных систем реального

времени, позволяющих моделировать взаимодействие команд агентов в динамическом, непрерывно меняющемся мире, с возможностью масштабирования на распределённые вычислительные ресурсы, поддержкой коммуникации между агентами и моделирования высокоуровневой сенсорной информации с шумами.

Для достижения указанной цели необходимо решить следующие задачи исследования:

1. Определить требования к функциональности среды имитационного моделирования многоагентных систем реального времени.
2. Провести анализ существующих среды имитационного моделирования многоагентных систем реального времени и выявить их ограничения в контексте определенных требований.

Требования к среде имитационного моделирования многоагентных систем реального времени

В работе [1] определены требования к СИМ МАС РВ. Среда должна позволять разработчику моделировать разные типы ИА, как в части робототехнической платформы (платформа для наземного беспилотника, для воздушного, для водного и др.), так и в части архитектуры интеллектуальной подсистемы агента. Среда должна позволять разработчику также добавлять новые компоненты для робототехнической платформы ИА (например, новые сенсоры). Еще одним требованием к среде является масштабируемость. Среда должна иметь возможность наращивать количество моделируемых объектов (как добавление ИА, так и сторонних объектов, участвующих в моделировании) без снижения общей производительности системы.

Главное требование к СИМ МАС РВ - моделирование в масштабе времени (задается тактом моделирования), который соответствует динамике процессов в моделируемой физической среде (именно это понимается под моделированием в реальном временем). Это требование важно, поскольку ИА, действующие в открытом, динамически меняющемся мире, представляются как системы ограниченной рациональности. Вследствие этого интеллектуальная подсистема агента должна выдавать команды управления физической подсистемой агента (команды движения платформы, отправки сообщения в канал межагентной связи, изменения параметров работы сенсоров и т.п.) в пределах такта моделирования среды, а СИМ МАС РВ должна обеспечить обработку этих команд.

Моделируемые сценарии командного взаимодействия и противодействия ИА требуют применения различных стратегий координаций действий в командах агентов. СИМ МАС РВ не должна накладывать ограничений на выбор стратегий, поскольку их выбор остается за разработчиком многоагентной системы. Среда должна лишь предоставить системные механизмы, позволяющие моделировать различные стратегии, в том числе как с применением коммуникации между ИА (с учетом пропускной способности канала, наличия шумов, дальности передачи), так и без ее применения.

В дополнение к моделированию каналов связи, выделяется также моделирование высокоуровневой сенсорной информации. Задача применения СИМ МАС РВ состоит в апробации интеллектуальных подсистем агентов, что не требует низкоуровневой сенсорной информации. Обработка низкоуровневой сенсорики приводит дополнительно к рассмотрению алгоритмов выделения высокоуровневой сенсорной информации и не является задачей разработчика МАС. По этой же причине не требуется использование высокоточного моделирования физической динамики и визуализации объектов среды.

Анализ существующих средств имитационного моделирования многоагентных систем реального времени

Для сравнения были выбраны 4 среды имитационного моделирования многоагентных систем: Gazebo, Robocup Soccer, MAgents, Unity ML Agents.

Gazebo [2] демонстрирует частичное соответствие требованиям к (СИМ МАС РВ). Совместно с Robot Operating System (ROS2), среда широко применяется для симуляции поведения разнородных агентов (воздушные, наземные и др.) в условиях приближённых к реальности. Gazebo обеспечивает высокоточную физику взаимодействия объектов, поддержку трёхмерной графики и модели низкоуровневых сенсоров (RGB-камеры, лидары и др.). Среда позволяют пользователю реализовать самостоятельно сенсоры для получения ВСИ и модули для моделирования межагентной коммуникации через добавления плагинов в Gazebo. Интеграция программных модулей роботов осуществляется с помощью инструмента. В ROS2 программные модули обмениваются информацией с помощью системы публикаторов/подписчиков. Пользователь может задавать несколько роботов в среде, но при этом возрастает вычислительная нагрузка, а также на каналы передачи данных между ROS2-модулями (приводит к задержкам получения данных [6]). Логика команд агентов не реализуется внутри среды. ИА и симулятор могут размещаться на разных вычислительных устройствах, при условии корректной настройки локальной сети ROS2. По умолчании среда моделирует агента с дискретным временем (такт моделирования задается в настройках), получая и отправляя сенсорную информацию согласно настройке тактов на публикаторах. Контроль над ходом симуляции возможен средствами самого Gazebo.

MAgent [3] – решение, частично удовлетворяющее требованиям. Представляет собой среду для моделирования поведения большого количества агентов на сетке (2D Grid World). Основное назначение — исследование многоагентных сценариев противодействия и кооперации с упором на масштабируемость. Время симуляции моделируется дискретно. Используется упрощенная физика на плоскости (перемещение клетка-клетка), доступна минимальная визуализация в виде точек на сетке. Среда позволяет моделировать десятки тысяч агентов в реальном времени, но в ограниченных сценариях на 2D-сетке. Сенсорная система ограничена фиксированным типом наблюдений, привязанных к клеткам сетки. Симуляция возможна только в 2D-пространстве. Нет поддержки абстрактных или настраиваемых сенсоров. Команды агентов задаются на уровне среды, что ограничивает разработчика МАС. Коммуникация между агентами не предусмотрена, а реализация затрудняется привязкой наблюдений к клеткам сетки.

RoboCup Soccer Simulator [4] еще один инструмент, частично удовлетворяющий требованиям. Среда ориентирована исключительно на моделирование футбольных матчей с участием автономных ИА. Среда имитирует поле и поведение игроков в 2D-пространстве (есть версии для 3D). Агенты функционируют в двух командах, получают сенсорную информацию и могут передавать сообщения друг другу. Каждый ИА может быть запущен на отдельной машине, что позволяет увеличивать вычислительные ресурсы, а подключение к серверу моделирования происходит по UDP-протоколу. Сенсорная информация и сообщения реализуются через строки, что требует дополнительной обработки и не масштабируется для сложных взаимодействий. Среда формирует команды агентов перед игрой при подключении клиента к серверу. Сенсорная модель фиксирована и привязана к футбольной тематике. Моделирование сценария производится дискретно с заданным

тактом времени, в ходе которого агенты должны получить и обработать сенсорную информацию от сервера, принять решение и отправить свое действие обратно серверу (удовлетворяет главному требованию реального времени СИМ МАС РВ).

Инструмент Unity ML-Agents [5] частично удовлетворяет требованиям СИМ МАС РВ. Используется игровой движок Unity как сервер моделирования, предоставляя возможности для настройки визуализации, управления объектами и логикой симуляции. Агенты запускаются на Python-клиентах и взаимодействую с Unity-сервером. В инструмент интегрированы библиотеки обучения интеллекта агента с помощью RL-алгоритмов. Среда фиксирует команды через параметр, задаваемый в конфигурации агента. Среда допускает написание собственной логики сенсоров и модуля канала межагентной связи. Однако сенсорные данные в среде представляются лишь в виде float-массивов (размерность в ходе симуляции фиксируется), изображений или лучевых сенсоров (RaySensor), что требует преобразования нестандартной информации в допустимый формат. Использование игрового движка в основе среды моделирования интеллектуальных агентов позволяет пользователю создавать сложную местность средствами Unity и не ограничивает в реализации различных типов агентов. Моделирование времени в среде дискретно, а в случае опоздания ответа агента среда останавливает симуляции до прихода ответа.

Заключение

Проведенный анализ существующих СИМ МАС показал, что ни одно из рассмотренных решений в полной мере не удовлетворяет всем требованиям к СИМ МАС РВ. Каждая из сред обладает сильными сторонами, но также имеет существенные ограничения.

Таким образом, для полноценного моделирования МАС в реальном времени требуется либо модификация существующих решений (что требует внесение модификаций в условиях архитектур инструментов), либо разработка специализированной среды, удовлетворяющей поставленным требованиям.

Дальнейшие исследования целесообразно направить на создание модульной СИМ МАС РВ, объединяющей преимущества рассмотренных инструментов и устраняющей их ключевые ограничения.

Список литературы

1. Пантелеев М.Г., Кохтенко Н.В., Лебедев С.В. Среда имитационного моделирования агентных систем реального времени // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 1 (77). – С. 54-57.

2. Gazebo Simulator [Электронный ресурс]. – Режим доступа: https://gazebosim.org/home (дата обращения: 02.04.2025).

3. Zheng L., Yang J., Cai H., Zhang W., Wang J., Yu Y. MAgent: A Many-Agent Reinforcement Learning Platform for Artificial Collective Intelligence // Proceedings of the AAAI Conference on Artificial Intelligence. - 2018. - Vol. 32, No. 1. (дата обращения: 05.04.2025).

4. RoboCup Soccer Simulator [Электронный ресурс]. – Режим доступа: https://rcsoccersim.github.io/ (дата обращения: 10.04.2025).

5. Juliani A., Berges V.-P., Teng E., Cohen A., Harper J., Elion C., Goy C., Gao Y., Henry H., Mattar M., Lange D. Unity: A General Platform for Intelligent Agents [Электронный ресурс] // arXiv. – 2020. – Режим доступа: https://arxiv.org/abs/1809.02627 (дата обращения: 14.04.2025). – arXiv:1809.02627 [cs.LG].

6. Kronauer T., Pohlmann J., Matthe M., Smejkal T., Fettweis G. Latency Analysis of ROS2 Multi-Node Systems [Электронный ресурс] // arXiv. - 2021. - Режим доступа: https://arxiv.org/abs/2101.02074 (дата обращения: 10.04.2025). - arXiv:2101.02074 [cs.RO].

ТЕХНОЛОГИЯ ОЦЕНКИ КАЧЕСТВА ЭЛЕКТРОТЕХНИЧЕСКИХ ИЗДЕЛИЙ МЕТОДАМИ МАШИННОГО ЗРЕНИЯ

М. А. Михайлова

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Аннотация. В данной работе рассмотрена задача визуальной оценки качества деталей электротехнических изделий. В статье приводится описание технологии обработки изображений деталей для решения задачи детектирования дефектов и их классификации, а также описание алгоритма позиционирования 3D-моделей изделий на основе их 2D-изображений.

Ключевые слова: оценка качества изделий, неразрушающий контроль, машинное зрение

Введение

В современном мире все чаще можно встретить использование электротехнических изделий в различных сферах жизни. Однако с увеличением объемов производства, растет проблема нехватки квалифицированных кадров. В то же время стремительное развитие научно-технических решений, особенно в области машинного обучения и нейронных сетей, способствует более широкому и качественному внедрению автоматизации процессов. Применение машинного зрения значительно упростит производственный процесс, в частности, решит задачу оценки качества изделий.

В рамках настоящей работы необходимо было предложить решение для автоматизации процесса выявления визуальных дефектов на электротехнических деталях малого размера. Решение должно корректно работать на деталях различной формы и материала, а также эффективно распознавать дефекты, предоставляя возможность дообучения в процессе использования.

В нынешних условиях данный этап производства исполняется группой операторов вручную, что является трудоемким, сложно масштабируемым и подверженным человеческому фактору решением.

При переходе к автоматическому контролю качества важно учитывать основные факторы, осложняющие работу, которые включают в себя сложность позиционирования и обеспечения однообразия освещения и фона при потоковой обработке мелких деталей.

Гипотеза: при корректном подборе инструмента и условий съемки, способов предварительной обработки изображений и методов машинного зрения возможна полная автоматизация процесса оптической оценки качества электротехнических деталей, включая обнаружение и классификацию дефектов.

В настоящее время рынок систем визуального контроля качества довольно обширный. В процессе анализа были рассмотрены несколько промышленных решений, представленных в продаже. Так, XCT8500 Universal Industrial X-ray Inspection System от компании Seamark ZM Technology может выполнять 2D, 3D и КТ для контроля качества, трехмерных измерений и неразрушающего анализа изделий. Другой пример, Sipotek Automated Visual Inspection Machinery, который предназначен для визуальной проверки мелких деталей, таких как винты, гайки и штампованные элементы. Система оснащена промышленными CCD-камерами с высоким разрешением и ПО для обнаружения внешних дефектов.

Несмотря на высокие технические характеристики, существующие решения имеют следующие проблемы: высокая стоимость оборудования и трудности логистики и последующей технической поддержки.

Учитывая вышеуказанные ограничения, было принято решение о разработке технологии, ориентированной на автоматизацию процесса выявления визуальных дефектов с учетом конкретных требований заказчика и ограничений производственной среды.

В докладе рассматриваются частные задачи обработки изображений в контексте целевой технологии контроля качества, предусматривающие использования модулей: удаления фона (УФ), поиска аномалий (ПА), добавления эталона (ДЭ).

Модуль удаления фона (УФ)

Удаление фона с изображения - процесс сепарации объекта от фона с целью минимизации отвлекающих факторов, которые могут находиться на фоне.

Стоит учитывать, что данный этап призван быть вспомогательным для корректной работы последующих модулей, поэтому основной задачей было найти решение, которое будет не только качественным с точки зрения изоляции объекта, но и малозатратным по времени. В процессе рассмотрения возможных методов решения стало ясно, что методы, основанные на нейронных сетях, несмотря на высокое качество сегментации, слишком чувствителен к бликам и отражениям, принимая их за объекты переднего плана. Кроме того, классические алгоритмы типа GrabCut [1], хотя и не требуют обучения, предполагают индивидуальный подбор параметров. Такой подход является избыточным в условиях, когда предполагается автоматическая обработка большого количества изображений с однотипными условиями съемки.

Учитывая специфику производственной среды, позволяющей заранее определить фон и освещение, логичным решением является удаление фона с помощью изначально заданного диапазона цветов. Такой способ показывает высокую скорость и эффективность в условиях контролируемой среды: с однородным фоном, стабильным освещением и центральным положением объекта.

Модуль поиска аномалий (ПА)

Модуль поиска аномалий нацелен на решение двух последовательных задач - детектирование и классификация дефектов. При детектировании аномалиями считаются любые отклонения от эталона, визуально заметные на поверхности детали. Эталон - идеальная деталь, не содержащая/содержащая допустимый минимум аномалий.

Задачей этапа предварительной обработки изображений является максимальное выражение дефектов при минимизации возможных бликов на металлических поверхностях, что достигается путем выравнивание гистограммы, обработки изображения с помощью размытия по Гауссу и использования высокочастотных фильтров.

Для выявления возможных дефектов используется два подхода: пороговый анализ разности пикселей и вычитание изображений с последующим сравнением по SSIM [2] (Structural Similarity Index).

Первый подход предполагает выявление участков, яркость которых значительно отличается от фона. Для этого используется разбиение изображения на регионы и проверка превышения заданного порога.

Некоторые области, отмеченные таким методом как аномалии, на деле являются частью детали. По этой причине важно обращать внимание на эталонное изображение. Сравнение

с эталонным изображением осуществляется через вычитание и метрику SSIM. Если сходство между соответствующими участками ниже порога, считаем, что участок содержит дефект.

Результаты работы двух методов - маски - объединяются с помощью логической операции И. Чтобы избежать разрыва дефектов, был реализован метод объединения контуров. Каждый контур анализируется, и для него рассчитываются основные характеристики, описывающие, насколько два контура могут быть частью одного и того же дефекта.

Для классификации был выбран ансамблевый подход. В случае использования ансамбля есть возможность подобрать алгоритмы так, чтобы выделить сильные стороны нескольких моделей и компенсировать недостатки за счет полученной устойчивости. Также, ансамблевые методы позволяют сделать архитектуру более адаптивной и удобной для долгосрочного применения в условиях ограниченного и постепенно расширяющегося датасета.

За основной ансамблевый алгоритм был выбран стекинг. В качестве базовых моделей взяты алгоритмы случайного леса, градиентного бустинга и K-ближайших соседей. Случайный лес и градиентный бустинг дополняют ансамбль, сочетая устойчивость к переобучению и способность последовательно устранять ошибки, а модель K-ближайших соседей усиливает его за счет учета локальных зависимостей, недоступных деревьям решений.

В качестве метамодели используется логическая регрессия с методом One-vs-Rest [3], который указывает на то, что для каждого класса создается свой бинарный классификатор и класс с наибольшей вероятностью считается результирующим.

Модуль добавления эталона (ДЭ)

Эталонным изображением может являться как готовое 2D-изображение детали с необходимого ракурса, так и 3D-модель в формате .stl. При использовании 3D-модели перед сравнением изображения дефектной детали с эталоном необходимо получить эталон в формате изображения.

Основным этапом является вращение объекта вокруг трех осей координат: X, Y и Z. Для каждой оси генерируется отдельная матрица вращения, а затем все они объединяются в единую матрицу преобразования, которая применяется к модели (см. рис. 1–2).

$$R_x(\alpha)=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\alpha & -\sin\alpha\\ 0 & \sin\alpha & \cos\alpha\end{pmatrix};\ R_y(\beta)=\begin{pmatrix}\cos\beta & 0 & \sin\beta\\ 0 & 1 & 0\\ -\sin\beta & 0 & \cos\beta\end{pmatrix};\ R_z(\gamma)=\begin{pmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{pmatrix} \quad (1)$$

$$R_{общ}(\alpha,\beta,\gamma)=R_x(\alpha)\,R_y(\beta)\,R_z(\gamma)$$

Рис.2 Единая матрица преобразования

После поворота создается сцена рендеринга: в нее добавляется сама модель, источник света, камера установлена под прямым углом и всегда направлена на центр модели.

Следующим шагом алгоритма является сравнение полученного изображения с эталонным изображением. Бинарная маска целевого изображения вычитается из маски модели и разница между пикселями двух изображений суммируется. Степень совпадения определяется по

следующему правилу: $S \to 0 \Leftrightarrow M_1 = M_2$, где S – итоговая сумма, M1 и M2 – маски сравниваемых изображений.

После выполнения поворота модели следующим этапом становится итеративный процесс уточнения ее позиции в плоскости XY. В каждой итерации, вокруг текущего положения равномерно размещаются точки, в каждой из которых вычисляется разница между полученным изображением и целевым (см. рис. 3). Ключевая особенность алгоритма — это динамическое уменьшение шага поиска. Как только удается достичь минимальной разницы между изображением модели и эталонным, алгоритм завершает работу. Итогом становится изображение 3D-модели, точно соответствующее целевому изображению.

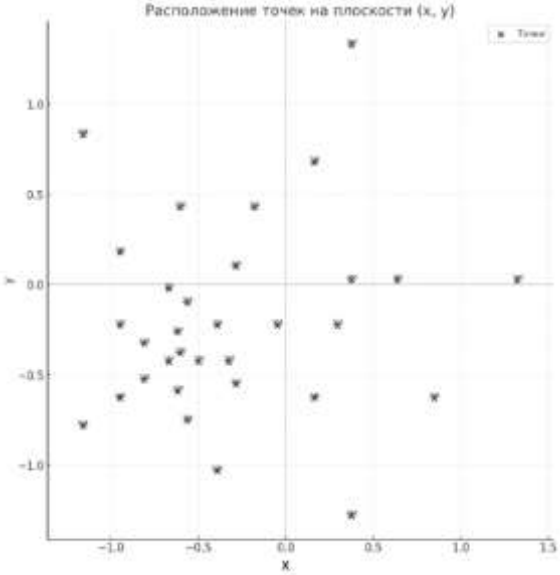


Рис.3 Пример расположения точек, рассматриваемых во время итеративного уточнения позиции

Результаты

В рамках проведенной работы было разработано модульное решение для автоматизации процесса выявления визуальных дефектов на электротехнических деталях малого размера с учетом ограничений производственной среды.

Для каждого из заявленных функциональных модулей подобран стек алгоритмов. Определены ограничения и рекомендации для снижения стоимости и повышения качества функционирования разработанной технологии.

С учетом свойств выбранных алгоритмов можно сделать вывод, что система будет демонстрировать удовлетворяющий уровень точности при корректной настройке параметров обработки и классификации. Вероятность ошибок в такой архитектуре может быть снижена за счет комбинирования методов анализа и ансамблевой классификации. Применимость предложенной технологии обусловлена ее гибкостью и адаптивностью к различным типам деталей, что делает ее перспективной для внедрения в промышленную среду.

Список литературы

1. Rother C., Kolmogorov V., Blake A. "GrabCut": interactive foreground extraction using iterated graph cuts // ACM Transactions on Graphics. — 2004. — №. Volume 23, Issue 3. — С. 309-314.
2. Wang Zh., Bovik A. C., Sheikh H. R., Simoncelli E. P. Image Quality Assessment: From Error Visibility to Structural Similarity // IEEE TRANSACTIONS ON IMAGE PROCESSING. — 2004. — №. VOL. 13, NO. 4. — С. 600–612.

3. *Платонов Е. Н., Мартынова И. Р.* Семантический анализ отзывов об организациях методами машинного обучения // Моделирование и анализ данных. 2024. Том 14. № 1. С. 7–26. DOI: 10.17759/mda.2024140101

МНОГОУРОВНЕВЫЙ ОБЪЯСНИТЕЛЬНЫЙ ИИ ДЛЯ ДИАГНОСТИКИ ПНЕВМОНИИ В РЕНТГЕН-СНИМКАХ: ОТ GRAD-CAM И LIME ДО ГЕНЕРАТИВНОГО АВТООПИСАНИЯ С MLOPS-ИНТЕГРАЦИЕЙ

Ю.В. Трофимов, И.П. Муравьев, Е.М. Кузнецов

Федеральное государственное бюджетное образовательное учреждение высшего образования «Университет «Дубна»

Аннотация. В статье рассматривается многоуровневый подход к объяснительному искусственному интеллекту для диагностики пневмонии в рентгенографии грудной клетки. Использование глубоких нейронных сетей и компьютерного зрения в медицине требует высокой точности и объяснимости результатов для специалистов. В рамках исследования интегрированы методы объяснительного ИИ: для локальной интерпретации - Grad-CAM, LIME, для глобальной интерпретации - нейро-нечёткая система ANFIS, что позволяет визуализировать важные области на снимках и формировать продукционные правила. Также разработан модуль генерации текстовых заключений. Разработка ориентирована на интеграцию в клинические системы, с возможностью масштабирования и гибкой настройки под различные сценарии применения. Работа направлена на повышение доверия к алгоритмам ИИ в медицине через улучшение их прозрачности.

Ключевые слова: Объяснительный искусственный интеллект, глубокие нейронные сети, компьютерное зрение, Grad-CAM, LIME, нейро-нечёткая система, генерация текстовых описаний, классификация пневмонии.

Введение

Использование глубоких нейронных сетей (Deep Learning, DL) и компьютерного зрения (Computer Vision, CV) для анализа рентгеновских (X-ray) снимков грудной клетки относится к активно развивающимся направлениям в области медицинской диагностики. При этом одного лишь высокого уровня точности алгоритмов, достигнутого за счёт современных CNN-моделей, оказывается недостаточно для их широкого применения на практике. Врачи и специалисты по лучевой диагностике стремятся видеть «прозрачную» логику работы системы и понимать, на чём основано итоговое решение. Именно по этой причине возрастающий интерес вызывает парадигма Explainable AI (XAI), которая помогает раскрыть механизм принятия решения нейросетью и сделать его интерпретацию наглядной и осмысленной [1].

В данной работе представлено расширение классических пост-хок XAI-подходов (Grad-CAM, LIME) за счёт интеграции нейронечёткого модуля для формирования продукционных правил, а также дополнение всего комплекса механизмом автоматической генерации диагностических описаний на естественном языке. Такое решение призвано иллюстрировать «горячие зоны» или ключевые факторы, влияющие на классификацию, а также повышает объяснимость за счёт наглядных логических правил и обеспечивает удобную для врача форму итогового вывода. Механизмы генерации текстового заключения учитывают результаты визуальных и логических объяснений и адаптированы для интеграции в клинические системы, будь то веб-приложения или PACS/EMR. Таким образом, создаётся более эффективный канал взаимодействия между алгоритмом и

пользователем, что потенциально повышает уровень доверия к моделям искусственного интеллекта и облегчает их внедрение в медицинскую практику.

Основная часть

В ходе исследования вначале был проведён подробный анализ нескольких общедоступных наборов рентген-снимков грудной клетки, включая Kaggle Pneumonia Dataset [2], NIH ChestXRay [3], MIMIC-CXR [4] и другие. Цель состояла в том, чтобы найти наиболее репрезентативный датасет, содержащий как бактериальные, так и вирусные случаи пневмонии. С учётом различных критериев (объём выборки, разнообразие патологий, качество маркировки/аннотаций) в итоге была выбрана коллекция Kaggle Chest X-Ray Images (Pneumonia) в качестве основной. Дополнительно сформировали вспомогательный набор из 50 снимков для ручной валидации и отладки объясняющих методов.

На втором этапе исследовались и настраивались (несколько предобученных архитектур свёрточных нейронных сетей (CNN), включая ResNet, VGG, DenseNet и EfficientNet, а также более компактные варианты вроде SqueezeNet, с использованием фреймворков PyTorch и TensorFlow. Для оценки качества моделей была применена k-fold кросс-валидация с вычислением метрик Accuracy, Precision, Recall, F1. Это позволило исключить менее многообещающие модели и выделить две наиболее стабильные по совокупности «качество–производительность–надёжность»: ResNet34 и DenseNet121. Итоговый выбор пал на ResNet34 как на модель, подлежащую дальнейшей интеграции в создаваемую XAI-систему.

На третьем этапе для повышения обобщающей способности (generalization) применились различные методы дополнения данных (data augmentation) (случайные повороты, отражения, изменение контрастности/яркости) и приёмы регуляризации (dropout, weight decay). Кроме того, был реализован инкрементный подход к классификации: сначала модель определяет факт наличия или отсутствия пневмонии (пороговое разделение «пневмония/норма»), а затем, при положительном ответе, уточняет её тип (вирусная или бактериальная). На данном шаге были получены результаты с точностью в диапазоне 90-94%, однако основным вызовом оставалось наглядное обоснование решений модели.

На четвёртом этапе была проведена интеграция многоуровневых методов объяснительного искусственного интеллекта (ОИИ), включающих Grad-CAM (Gradient-weighted Class Activation Mapping), LIME (Local Interpretable Model-Agnostic Explanations) и Adaptive Neuro-Fuzzy Inference System (ANFIS). Подобное комбинирование обеспечивает комплексную интерпретацию результатов на нескольких уровнях:

- Локальный уровень (Grad-CAM, LIME): выделяются наиболее значимые зоны на рентгеновском изображении и формируются локальные суррогатные модели, позволяющие понять, какие признаки повлияли на решение нейросети в конкретном случае;
- Глобальный уровень (ANFIS): нейро-нечёткая система объединяет обучаемую нейронную компоненту с логикой нечёткого вывода, что не только снижает чувствительность модели к шумам и обеспечивает стабильность работы, но и даёт возможность преобразовать внутренние представления сети в наглядные продукционные правила [5,6].

Таким образом, Grad-CAM и LIME способствуют детальному анализу на уровне единичного снимка, тогда как ANFIS придаёт прозрачность всей структуре решений,

формируя мост между «чёрным ящиком» нейронных весов и интерпретируемыми логическими посылками. Это облегчает верификацию работы алгоритма, понижает вероятность ошибки и искажения выходных данных, что в совокупности повышает уровень доверия к модели со стороны врачей.

На пятом, предпоследнем этапе был разработан высокоуровневый модуль автоматизированной генерации текстовых заключений, позволяющий объединить результаты локальных (Grad-CAM, LIME) и глобальных (ANFIS) интерпретаций в единый когерентный контекст. Данная система извлекает «горячие зоны» и ключевые признаки, выявленные инструментами XAI, и посредством «мягких дескрипторов» и набора семантических шаблонов в области медицины (на базе предобученных языковых моделей) формирует связное диагностическое описание.

В результате специалисту предоставляется лаконичный, но информативный текст, подтверждающий или опровергающий факт патологии, что выгодно дополняет привычную визуализацию. Подобная многостадийная архитектура (CNN-классификация → многоуровневая XAI-интерпретация → генеративное автоописание) предлагает одновременно наглядность и доступное объяснение, позволяя врачу быстро сформировать целостное представление о работе модели. Такой ориентированный на пользователя подход (human-centric XAI) обеспечивает необходимый уровень прозрачности и доверия со стороны медицинских специалистов, что является критически важным фактором для успешного внедрения алгоритмов искусственного интеллекта в клиническую практику.

На заключительном этапе была реализована многоуровневая веб-платформа, ориентированная на интеграцию в сложные клинические контуры (PACS/EMR). Архитектура решения построена с учётом принципов микросервисности и контейнеризации, что обеспечивает масштабируемость и возможность гибкой настройки под различные сценарии использования. В результате пользователь (врач) при загрузке рентген-снимка взаимодействует с единой средой, получая:

- Онтологически-прозрачную классификацию: бинарное решение (наличие или отсутствие пневмонии), основанное на свёрточной нейронной сети и нейро-нечётком модуле.
- Интерпретируемые визуальные пояснения: метрики активации (Grad-CAM, LIME) и генерируемые «горячие зоны» для анализа локальных участков.
- Глобальные логические правила: продукционные сценарии, формируемые нейро-нечётким механизмом (NFIS), повышающие прозрачность выводимых результатов.
- Текстовое автоописание: автоматизированные диагностические заключения на естественном языке, синтезируемые на основе предобученных языковых моделей и медико-семантических шаблонов.

Для повышения доверительности и надёжности система может быть интегрирована в контур MLOps (CI/CD-пайплайны, мониторинг метрик и версионность моделей), что упрощает поддержку решения и обновление весов нейросетей в условиях реальной клинической эксплуатации. Предусмотрены интерфейсы (REST/HL7/FHIR) для подключения к существующим EMR/PACS, а также набор инструментов для Clinical Decision Support (CDS), позволяющий автоматически формировать подсказки при анализе рутинных протоколов. Такая комплексная, ориентированная на пользователя XAI-платформа отвечает строгим требованиям к объяснимости, точности и скорости,

необходимым для успешного внедрения алгоритмов искусственного интеллекта в современную медицинскую практику.

Выводы

Представленный многоуровневый XAI-подход, включающий двухфазную классификацию (основная детекция + уточнение подтипа пневмонии), визуальные интерпретации (Grad-CAM, LIME) и механизмы генеративной репрезентации (автоматическое формирование текстовых описаний), демонстрирует высокие показатели точности, а так же существенный рост прозрачности при анализе рентгеновских снимков грудной клетки. Благодаря интеграции локальных (Grad-CAM, LIME) и глобальных (нейро-нечёткая система) стратегий интерпретации, а также автоописаний на естественном языке, система формирует интуитивно понятный «мост» между глубоким обучением и клиническими рабочими процессами.

Разработанное веб-приложение, ориентированное на встраивание в PACS/EMR, обеспечивает единый интерфейс, где врач может оперативно оценить диагностические результаты, визуальные тепловые карты и текстовые заключения. Такая архитектура, сочетающая методы ОИИ и удобные прикладные интерфейсы, облегчает MLOps-ориентированную интеграцию в инфраструктуру медицинских учреждений, повышает доверие к системе и способствует более широкому внедрению алгоритмов искусственного интеллекта в практику рентгенодиагностики.

Список литературы

6. E. Tjoa and C. Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 11, pp. 4793-4813, Nov. 2021, doi: 10.1109/TNNLS.2020.3027314.
7. Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), “Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification”, Mendeley Data, V2, doi: 10.17632/rscbjbr9sj.2
8. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri and R. M. Summers, "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 3462-3471, doi: 10.1109/CVPR.2017.369.
9. Johnson, A.E.W., Pollard, T.J., Berkowitz, S.J. *et al.* MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Sci Data* **6**, 317 (2019). https://doi.org/10.1038/s41597-019-0322-0
10. Jang, S. J. F. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics, 23*(3), 665–685 1.
11. Averkin, A. N. Possibilities of Application of Neuro-Fuzzy Networks for Ophthalmologic Image Classification / A. N. Averkin, E. N. Volkov, S. A. Yarushev // *Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications*. – 2024. – Vol. 34, No. 3. – P. 610-616. – doi: 10.1134/S1054661824700421. – EDN IMZIJT.

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № 124112200072-2).

АНАЛИЗ МЕТОДОВ ПОИСКА ПУТИ ПОДВИЖНОГО АГЕНТА В ДИНАМИЧЕСКОЙ СРЕДЕ С ИСПОЛЬЗОВАНИЕМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Новицкий М.Д., Грибов А.В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Аннотация. В статье рассмотрены различные алгоритмы поиска пути подвижного агента в динамической среде с использованием обучения с подкреплением по следующим критериям: тип пространства, энергетические ограничения, количество динамических препятствий. В результате анализа показано, что большинство алгоритмов решают задачу в трёхмерном пространстве, но при этом без учёта большинства факторов, влияющих на заряд батареи и с небольшим количеством динамических препятствий.

Ключевые слова: поиск пути подвижного агента, динамическая среда, обучение с подкреплением

Введение

В наше время стремительно развиваются сферы применений подвижных агентов, как в России так и в мире в целом [1, 2]. Данный рост приводит к увеличению количества предлагаемых алгоритмов, связанных с проблемой поиска пути подвижного агента в динамической среде. Из-за роста числа предлагаемых алгоритмов появляется необходимость в выявлении их слабых и сильных сторон.

Целью данной работы является анализ существующих алгоритмов поиска пути для подвижного агента в динамической среде, основанных на обучении с подкреплением, и выявление критериев, предъявляемых к таким алгоритмам. Исходя из цели статьи были поставлены следующие задачи:

1. Сделать обзор и сравнительный анализ основных алгоритмов поиска пути для подвижного агента в динамической среде, основанных на обучении с подкреплением.
2. Выявить тенденции развития существующих алгоритмов поиска пути для подвижного агента в динамической среде, основанных на обучении с подкреплением.
3. Определить критерии, предъявляемые к алгоритмам поиска пути для подвижного агента в динамической среде, основанные на обучении с подкреплением.

Принцип отбора аналогов

В качестве аналогов были рассмотрены алгоритмы планирования пути подвижного агента в динамической среде, основанные на обучении с подкреплением. Основным критерием, предъявляемым к выбору аналогов являлась успешность решения задачи планирования пути подвижного агента в динамической среде. Подбор аналогов осуществлялся с помощью Google Scholar и поисковой системы Google. Для поиска аналогов применялись следующие запросы: "Reinforcement learning models for UAV route construction in a dynamic environment", "UAV path planning reinforcement learning algorithm in a dynamic environment", "Algorithm for drone path panning in dynamic environment", "UAV path planning reinforcement learning methods".

Yan C. с соавторами предлагают вариант решения планирования пути движения подвижного агента, в котором агент состоит из 3 модулей: модуль оценки ситуации, модуль Dueling Double Deep Q-networks (D3QN) и модуль выбора действия [3]. Модуль оценки ситуации на основе полученных данных строит карту ситуации. Карта ситуации показывает обзор тактической ситуации в определенное время, что помогает принимать тактические решения. После этого модуль D3QN использует стек карт ситуаций и предсказывает значения Q всех возможных действий. Модуль выбора действий на основе всех Q выбирает следующее действие. На основе этого выбора подвижный агент перемещается на новую позицию.

В статье [4] рассмотрен метод, в котором задача поиска пути разбивается на три простых подзадачи: избегание объекта, достижения цели и принятия решения. Таким

образом задачи избегания и достижения решаются с помощью DLR сети, которая разбита на две подсети: сеть избегания и сеть достижения. Задача принятия решения регулируется ~~с помощью~~ урезанной DLR функции потерь, которая должна выбрать определённое действие из сети избегания и сети достижения для получения награды.

В решении, предлагаемом в статье [5], при обучении с подкреплением используется алгоритм Proximal Policy Optimization (PPO). PPO обучается с использованием оценки состояния (value function) для стабильности. Применяется ограничение на обновления политики для предотвращения нестабильных изменений.

В статье [6], предлагается метод, основанный на использовании рекуррентных сетей. Архитектура сети: Используется комбинация CNN для извлечения пространственных признаков и RNN для обработки временных последовательностей. А для принятия решения используется стратегия RQ, которая объединяет текущую награду R и текущую оценку-Q. Также предлагается алгоритм, который комбинирует механизм выбора решений искусственного потенциального поля с алгоритмом Dueling Double Deep Q-networks (D3QN) [7].

Критерии, предъявляемые к алгоритмам поиска пути подвижного агента

Основными характеристиками рассматриваемых методов являются тип пространства, энергетические ограничения и количество динамических препятствий.

Тип пространства показывает для какого типа пространства был разработан алгоритм поиска пути подвижного агента в динамической среде с использованием обучения с подкреплением. Алгоритмы могли быть разработаны для двухмерного или трёхмерного пространства.

Энергетические ограничения показывают какие факторы влияющие на заряд батареи, отличные от длины пути или столкновения с препятствиями, учитывались алгоритмом. Дополнительными факторами могут выступать такие вещи как: вес, ветер, тяга двигателя, маневры подвижного агента, которые могут повлиять на работу батареи.

Количество динамических препятствий показывает сколько динамических препятствий находилось в среде, в которой тестировался алгоритм.

Основные характеристики методов представлены в таблице 1.

Таблица 1

Основные характеристики методов

Метод	Тип пространства	Энергетические ограничения	Количество динамических препятствий
Towards Real-Time Path Planning through Deep Reinforcement Learning for a UAV in Dynamic Environments	трёхмерное пространство	Штраф за неаккуратные манёвры	3
UAV navigation in high dynamic environments: A deep reinforcement learning approach	двумерное пространство	Штраф за неаккуратные манёвры	12
Dynamic Navigation in Unconstrained Environments Using Reinforcement Learning Algorithms	трёхмерное пространство	Штраф за неаккуратные манёвры	не указано
Unmanned Aerial Vehicle Path Planning Algorithm Based on Deep Reinforcement Learning in Large-Scale and Dynamic Environments	трёхмерное пространство	Штраф за неаккуратные манёвры	4

Метод	Тип пространства	Энергетические ограничения	Количество динамических препятствий
Path Planning for Autonomous Vehicles in Unknown Dynamic Environment Based on Deep Reinforcement Learning	трёхмерное пространство	Штраф за неаккуратные манёвры	10

Как видно из приведенных в таблице данных большинство методов поиска пути подвижного агента на основе обучения с подкреплением решают задачу в трёхмерном пространстве, в котором находится лишь небольшое количество динамических препятствий. Однако, при этом алгоритмы учитывают лишь один фактор который влияет на заряд батареи - маневры подвижного агента. Данный недочёт приводит к более быстрому расходу заряда, что, в свою очередь, влияет на дальность полёта вплоть до невозможности завершить маршрут. Исходя из этого можно выделить алгоритм из статьи [7], так как он решает задачу в трёхмерном пространстве и по сравнению с другими алгоритмами тестировался на большем количестве препятствий, хотя он также игнорирует большинство факторов влияющих на заряд батареи.

В результате анализа было показано, что ни один из методов не учитывает большинство факторов, влияющих на заряд батареи. Поэтому–для решения проблемы поиска пути подвижного агента в динамической среде с использованием обучения с подкреплением, необходимо разработать метод, который будет учитывать следующие критерии: реализация планирования пути в трёхмерном пространстве, как минимум двадцать динамических препятствий и факторы, влияющие на заряд батареи такие как: вес, ветер, тяга двигателя, маневры подвижного агента.

Заключение

В ходе работы был проведен анализ существующих решений задачи поиска пути подвижного агента в динамической среде с помощью обучения с подкреплением. Были выявлены следующие критерии для анализа методов: тип пространства, энергетические ограничения, количество динамических препятствий. Направление дальнейших исследований - разработка метода поиска пути подвижного агента в динамической среде на основе обучения с подкреплением, учитывающего предложенные критерии, особенно факторы, влияющие на заряд батареи.

Список литературы

1. Unmanned aerial vehicles market size: Mordor Intelligence. Report // Industry Analysis, Size & Forecast. https://www.mordorintelligence.com/industry-reports/uav-market (дата обращения: 31.10.2024).
2. Grand View Research. Commercial UAV Market Analysis by Product (Fixed Wing, Rotary Blade, Nano, Hybrid), by Application (Agriculture, Energy, Government, Media & Entertainment) and Segment Forecasts to 2022; ID: 978-1-68038-584-7; Grand View Research: San Francisco, CA, USA; p. 117.
3. Yan C., Xiang X., Wang C. Towards real-time path planning through deep reinforcement learning for a UAV in dynamic environments //Journal of Intelligent & Robotic Systems. – 2020. – V. 98. – P. 297-309.
4. Tong G. U. O. et al. UAV navigation in high dynamic environments: A deep reinforcement learning approach //Chinese Journal of Aeronautics. – 2021. – V. 34. – №. 2. – P. 479-489.
5. Chronis C. et al. Dynamic navigation in unconstrained environments using reinforcement learning algorithms //IEEE Access. – 2023.
6. Xie R. et al. Unmanned aerial vehicle path planning algorithm based on deep reinforcement learning in large-scale and dynamic environments //IEEE Access. – 2021. – V. 9. – P. 24884-24900.
7. Hu H. et al. Path planning for autonomous vehicles in unknown dynamic environment based on deep reinforcement learning //Applied Sciences. – 2023. –V. 13. – №. 18. – P. 10056.

ПРЕИМУЩЕСТВА И НЕДОСТАТКИ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

Д. А. Прахов

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Современная медицина сталкивается с вызовами в диагностике заболеваний, что требует внедрения инновационных технологий. Нейронные сети, продемонстрировавшие свою эффективность в различных областях, становятся важным инструментом в медицинской диагностике. Данное исследование направлено на анализ их преимуществ и недостатков в этой сфере. Основные цели включают изучение повышения точности диагностики, ускорения обработки данных и снижения вероятности ошибок. Также рассматриваются ограничения, такие как необходимость больших объемов данных и сложности интерпретации результатов. Выбор нейронных сетей как объекта исследования обусловлен их потенциалом для трансформации медицинской практики. Результаты анализа помогут разработать рекомендации по эффективному внедрению этих технологий в медицину.

Ключевые слова: современная медицина, диагностика заболеваний, инновационные технологии, нейронные сети, эффективность, медицинская помощь, преимущества и недостатки, точность диагностики, обработка данных, ограничения и риски, сверточные нейронные сети (cnn), рекуррентные нейронные сети (rnn), глубокие нейронные сети (dnn), гибридные архитектуры, увеличение точности, ускорение анализа данных, снижение человеческой ошибки, большие объемы данных, интерпретация результатов, этические аспекты, правовые аспекты, инфраструктура, безопасность, эффективность, будущее нейронных сетей, перспективы развития, мультимодальные системы, объяснимый искусственный интеллект, стандартизация использования.

Нейронные сети представляют собой математические модели, вдохновленные структурой и работой биологических нейронов. Они состоят из множества слоев, содержащих искусственные нейроны, которые обрабатывают и передают информацию. В зависимости от структуры и назначения различают несколько типов нейронных сетей, каждая из которых обладает уникальными характеристиками. Эти особенности делают сети пригодными для решения различных задач, включая медицинскую диагностику. В медицине используются такие типы нейронных сетей, как сверточные, рекуррентные, глубокие и гибридные архитектуры. Введение в их основные принципы и применение является ключевым шагом для понимания их роли в диагностике.

Сверточные нейронные сети (CNN) обрабатывают двумерные данные, такие как изображения, и выделяют ключевые признаки, что делает их полезными для анализа медицинских снимков, включая рентген и МРТ. Их способность выявлять сложные паттерны повышает точность диагностики.

Рекуррентные нейронные сети (RNN) работают с последовательными данными, например, ЭКГ и мониторингом жизненно важных показателей. Они учитывают временные зависимости, что важно для предсказания и классификации.

Глубокие нейронные сети (DNN) имеют много слоев и обрабатывают сложные данные, применяясь в анализе геномной информации и предсказании заболеваний. Их способность работать с большими объемами данных делает их важными в медицине.

Гибридные архитектуры объединяют преимущества CNN и RNN, позволяя анализировать как изображения, так и временные ряды. Это открывает новые возможности

для точной диагностики, а использование эволюционного подхода усиливает эффективность таких моделей, адаптируя их к разнообразным данным и задачам.

Примеры успешного применения нейронных сетей в медицинской диагностике демонстрируют их эффективность и потенциал. Исследование, опубликованное в журнале 'Nature Medicine' в 2019 году, показало, что нейронные сети достигли точности 94,5% при диагностике рака кожи, что сопоставимо с результатами ведущих дерматологов. В 2020 году в Китае была разработана нейронная сеть для диагностики COVID-19 на основе анализа КТ-изображений, которая продемонстрировала точность 92% и значительно ускорила процесс диагностики. Эти примеры подчеркивают, как использование нейронных сетей может не только улучшить точность диагностики, но и ускорить процесс принятия медицинских решений. «Использование в медицинской диагностике методов распознавания и создание на их основе автоматизированных систем коренным образом изменяет эффективность диагностики с точки зрения широты и глубины охвата медицинских симптомов, ее оперативности, полноты и достоверности» (Кузнецова, Протченко, Боброва, 1 с.). Это особенно актуально в условиях критической нехватки времени.

Нейронные сети значительно превосходят традиционные методы в обработке больших объемов медицинских данных благодаря параллельным вычислениям и адаптации к сложным структурам. Исследования показывают, что «обработка изображений с использованием нейронных сетей позволяет значительно улучшить точность, скорость и качество обработки» (Курбансахатов, 2024, 40 с.). По данным журнала Nature Medicine, ИИ анализирует медицинские изображения в десять раз быстрее специалистов, сохраняя сопоставимую точность. Это особенно важно в условиях растущего объема медицинских данных, требующих эффективных инструментов для обработки.

Нейронные сети находят применение в экстренной медицине для оперативной диагностики. Например, в 2020 году в Великобритании использование данных технологий позволило сократить время анализа компьютерной томографии легких пациентов с COVID-19 с нескольких часов до нескольких минут.

Нейронные сети играют важную роль в снижении вероятности человеческой ошибки в медицинской диагностике. Благодаря способности анализировать огромные объемы данных и выявлять сложные взаимосвязи, которые могут быть упущены человеком. Согласно данным Всемирной организации здравоохранения, около 10% медицинских ошибок связаны с неправильной интерпретацией данных. Нейронные сети, обученные на большом количестве медицинских записей, способны минимизировать эти ошибки за счет автоматизации анализа и предоставления более точных результатов. Это особенно важно в сложных случаях, где человеческий фактор может сыграть решающую роль в постановке диагноза. Исследование 2020 года выявило, что использование нейронных сетей в диагностике рака кожи позволило сократить частоту ошибок на 30% по сравнению с традиционными методами. Подобные достижения подчеркивают значимость интеграции нейронных сетей в медицинскую практику, что способствует повышению качества диагностики и безопасности пациентов.

Современные нейронные сети, используемые в медицинской диагностике, требуют значительных объемов данных для обучения. Это обусловлено их архитектурой, базирующейся на большом количестве параметров, которые должны быть оптимизированы для выполнения конкретных задач. Например, для обучения сложных моделей, таких как GPT-3, требуется до 570 гигабайт текстовых данных. Однако в медицинской сфере сбор

достаточного объема данных может быть затруднен. Медицинские данные часто имеют неструктурированный характер, что затрудняет их использование для обучения нейронных сетей. Исследование IBM показывает, что около 80% медицинских данных представляют собой неструктурированные формы, такие как записи врачей, результаты обследований и изображения. Это создает преграды для их обработки и анализа. Кроме того, доступ к данным может быть ограничен из-за строгих требований к конфиденциальности и правовым ограничениям. Такие проблемы значительно замедляют разработку и внедрение нейронных сетей в медицинскую диагностику, требуя значительных усилий для преодоления этих барьеров.

Одной из ключевых проблем внедрения нейронных сетей в медицину является необходимость объяснения их решений специалистам. Нейронные сети, являясь сложными вычислительными моделями, часто воспринимаются как «черный ящик», что связано с трудностью интерпретации их внутренних процессов. В исследовании 2019 года было установлено, что лишь 15% врачей полностью доверяют выводам моделей ИИ, не понимая их механизмов. Это ограничивает применение нейронных сетей в медицинской практике, поскольку специалисты не могут в полной мере оценить корректность и надежность полученных выводов. Врачи и медицинский персонал должны быть уверены в точности и обоснованности диагноза, предложенного системой. Всемирная организация здравоохранения в 2021 году отметила важность разработки объяснимых моделей ИИ для повышения доверия и понимания среди медицинских работников. Это подчеркивает необходимость создания инструментов и подходов, которые сделают работу нейронных сетей более доступной для интерпретации специалистами. Развитие объяснимых моделей может существенно улучшить взаимодействие между медицинским персоналом и технологиями, способствуя более эффективному применению искусственного интеллекта в здравоохранении. На основании отчёта IBM Global AI Adoption Index 2024 выявлено, что наибольшие препятствия во внедрении нейронных сетей являются: конфиденциальность данных, проблемы доверия/прозрачности, отсутствие навыков для его реализации.

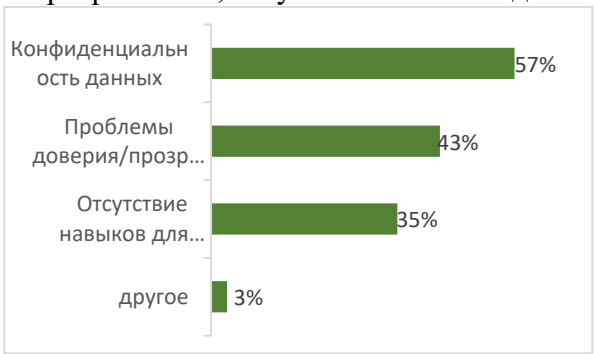


Рисунок 1. Результаты опроса: What are your organization's biggest inhibitors to adopting generative AI?

Это свидетельствует о важности создания правовых рамок, которые обеспечат безопасность и надежность технологий, а также защиту данных пациентов. Регулирование должно учитывать необходимость сертификации и проверки алгоритмов, чтобы гарантировать их эффективность и соответствие медицинским стандартам. Важно, чтобы эти законы не только защищали интересы пациентов, но и способствовали развитию инноваций в области медицины.

Ключевым фактором успешного внедрения нейронных сетей в медицинскую диагностику является наличие качественных и объемных данных. Для обучения моделей глубокого обучения требуется большое количество разнообразных данных, как показало исследование в Nature Medicine, где использовалось более 130 тысяч изображений для диагностики рака кожи. Важны не только объем и разнообразие данных, но и их качество, которое должно обеспечиваться тщательной аннотацией и проверкой экспертами. Кроме того, необходима соответствующая инфраструктура для обработки больших объемов данных и выполнения сложных вычислений. Однако, по данным ВОЗ, лишь 60% медицинских учреждений в странах с низким доходом имеют доступ к такой инфраструктуре, что создает препятствия для использования нейронных сетей. Для успешного внедрения технологий требуется создание и поддержание инфраструктуры, включая серверы и облачные платформы, для обеспечения доступа к современным технологиям на глобальном уровне.

Безопасность нейронных сетей в медицинской диагностике требует внимания. Инструменты интерпретации, такие как Grad-CAM, помогают врачам понять алгоритмы и повысить доверие к искусственному интеллекту, что способствует его безопасному внедрению. Исследование в Nature Medicine (2020) показало, что такие методы увеличивают доверие специалистов. Эффективность нейронных сетей возрастает благодаря предварительно обученным моделям и облачным вычислениям, что сокращает время обучения и улучшает обработку данных. По данным ВОЗ (2021), эти подходы повысили эффективность обработки медицинских изображений на 45%. Искусственный интеллект в здравоохранении обещает улучшение диагностики и прогнозирования заболеваний.

Точность диагностики нейронных сетей сопоставима с квалифицированными специалистами, например, алгоритмы для диагностики рака кожи достигают 95% точности. С развитием технологий нейронные сети могут стать более эффективными, ускоряя анализ данных и снижая вероятность ошибок. Улучшение алгоритмов интерпретации повысит доверие врачей и ускорит интеграцию в клиническую практику. Исследования направлены на разработку алгоритмов для диагностики редких заболеваний и внедрение искусственного интеллекта, что может сократить расходы на здравоохранение до 25%. Создание мультимодальных систем, объединяющих различные типы данных, также обещает улучшить качество диагностики и персонализацию лечения.

Список литературы

1. Cardoso M. J., Li W., Brown R., Ma N., Kerfoot E., Wang Y., Murrey B., Myronenko A., Zhao C., Yang D., Nath V., He Y., Xu Z., Hatamizadeh A., Zhu W., Liu Y., Zheng M., Tang Y., Yang I., Zephyr M., Hashemian B., Alle S., Zalbagi Darestani M., Budd C., Modat M., Vercauteren T., Wang G., Li Y., Hu Y., Fu Y., Gorman B., Johnson H., Genereaux B., Erdal B. S., Gupta V., Diaz-Pinto A., Dourson A., Maier-Hein L., Jaeger P. F., Baumgartner M., Kalpathy-Cramer J., Flores M., Kirby J., Cooper L. A. D., Roth H. R., Xu D., Bericat D., Floca R., Zhou S. K., Shuaib H., Farahani K., Maier-Hein K. H., Aylward S., Dogra P., Ourselin S., Feng A. MONAI: An open-source framework for deep learning in healthcare // Springer Nature. — 2021. — [б. и.].
2. Кузнецова О. В., Протченко Е. В., Боброва Т. С. Применение нейронных сетей для решения задач медицинской диагностики // [б. и.]. — [б. м.], [б. г.]. — [б. и.].
3. Научная инициатива: проблемы и перспективы внедрения инновационных решений: сборник статей Международной научно-практической конференции (27 ноября 2024 г, г. Магнитогорск). — Магнитогорск: МЦИИ ОМЕГА САЙНС, 2024. — 294 с.
4. Научные достижения и инновационные подходы: фундаментальные и прикладные аспекты. Сборник научных трудов по материалам I Международной научно-практической конференции (г.-к. Анапа, 25 декабря 2024 г.) / под ред. Скориковой Е. Н. — Анапа: Изд-во «НИЦ ЭСП» в ЮФО, 2024. — 96 с.

5. Сезнева Т.Н., Павлова В.Б., Ващалова Н.А. Опыт использования метода БОС по параметрам ЭМГ в восстановительном лечении детей с идиопатическим сколиозом // Медицина. Социология. Философия. Прикладные исследования. — 2022. — № 2. — С. 23-28.
6. Цой Ю. Р., Спицын В. Г. Эволюционный подход к настройке и обучению искусственных нейронных сетей // Нейроинформатика. — 2006. — Т. 1, № 1. — С. 34–35.
7. IBM GLOBAL AI ADOPTION INDEX – ENTERPRISE REPORT N O V E M B E R 8 – 2 3 , 2 0 2 3.

ПРИМЕНЕНИЕ 4D ГАУССОВСКОГО СПЛАТТИНГА ДЛЯ МОДЕЛИРОВАНИЯ ДИНАМИЧЕСКОЙ ТРЁХМЕРНОЙ СЦЕНЫ

А.М. Троицкая, Е.Д. Фабричева, Д.В. Чепасов

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В работе рассматривается современный метод моделирования сцен, изменяющихся во времени, – динамический гауссовский сплаттинг. Описываются принципы работы алгоритма, этапы оптимизации и подход к деформации 3D-гауссиан. Проводится исследование качества восстановления сцены для синтетических и реальных наборов данных по нескольким метрикам сравнения изображений.

Ключевые слова: 4D-сплаттинг, 3D-сплаттинг, реконструкция сцен

Введение

Моделирование трёхмерных сцен представляет значительный интерес в различных областях, включая робототехнику, компьютерное зрение, виртуальную и дополненную реальность, разработку видеоигр и создание визуальных эффектов.

Одним из ключевых достижений последних лет в этой области стал гауссовский сплаттинг — метод представления сцены в виде множества гауссиан, каждая из которых характеризуется координатами положения, ковариационной матрицей, цветом и альфа-каналом. По сравнению с методами семейства Neural Radiance Fields (NeRF) [11], гауссовский сплаттинг демонстрирует более высокую производительность за счёт явного и легко оптимизируемого представления, что сделало его популярным в исследованиях и индустрии.

Применение гауссовского сплаттинга к динамическим сценам вызывает ряд трудностей, связанных с необходимостью учёта временных изменений объектов. Наиболее очевидным подходом является использование статического сплаттинга для каждого кадра по отдельности. Однако это требует значительных вычислительных ресурсов и не обеспечивает временной согласованности между кадрами. Для преодоления ограничений были разработаны методы динамического, или 4D, гауссовского сплаттинга, которые моделируют деформацию гауссиан во времени и используют пространственно-временные зависимости для повышения эффективности и качества реконструкции.

В работе рассматривается метод 4D гауссовского сплаттинга, описанный в статье «4D Gaussian Splatting for Real-Time Dynamic Scene Rendering» [4] и его применение к тестовым данным.

Гауссовский сплаттинг

Гауссовский сплаттинг представляет собой эффективный подход к представлению и рендерингу трёхмерных сцен, который соответствует предыдущему state-of-the-art методу

NeRF в качестве сгенерированных изображений и превосходит его в скорости обучения [2], достигая синтеза видов сцены с новых ракурсов в реальном времени – более 30 кадров в секунду [1]. Входными данными служат фотографии с известными положениями камеры в пространстве и объёмные представления сцены, чаще всего облака точек [2]. Сцена моделируется как совокупность трёхмерных гауссиан, каждая из которых характеризуется следующими параметрами: центральная точка $\mu \in \mathbb{R}^3$, ковариационная матрица $\Sigma \in \mathbb{R}^{3\times3}$, цвет c, представленный через сферические гармоники, и прозрачность $\alpha \in [0,1]$. В качестве центров гауссиан μ выступают точки из облака. Каждая гауссиана может быть выражена как:

$$G(x) = e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}.$$

Ковариационная матрица может быть декомпозирована на масштабирование и поворот $\Sigma = RSS^T R^T$, где $R \in SO(3)$ — матрица вращения, $S \in \mathbb{R}^{3\times3}$ — диагональная матрица масштабирования.

Процесс обучения заключается в оптимизации путем градиентного спуска параметров гауссиан таким образом, чтобы сплаттированные изображения максимально соответствовали входным. Это делает возможным обучение всей системы без необходимости предварительной обработки и сложных сетевых архитектур.

Особенностью метода является адаптивное управление плотностью гауссиан во время оптимизации: если слишком маленькая гауссиана находится недостаточно хорошо реконструированном месте сцены, то её дублируют; если гауссиана слишком велика, то её заменяют двумя меньшими для улучшения детализации; гауссианы с высоким значением прозрачности удаляются.

Рендеринг гауссиан осуществляется с использованием быстрого тайлового рендерера, оптимизированного под GPU. Изображение делится на тайлы размером 16×16 пикселей, и каждый тайл обрабатывается параллельно. Алгоритм рендеринга включает проекцию гауссиан на плоскость камеры [3], сортировку по глубине и α-смешивание для каждого пикселя. Процесс α-смешивания выглядит следующим образом:

$$C_{\text{final}} = \sum_{i=1}^{N} \alpha_i \prod_{j=1}^{i-1} (1-\alpha_j) c_i, \quad \alpha_{\text{final}} = 1 - \prod_{i=1}^{N} (1-\alpha_i),$$

где c_i — цвет i-й гауссианы, α_i — её прозрачность.

Для работы с динамическими сценами в статье «4D Gaussian Splatting for Real-Time Dynamic Scene Rendering» был разработан метод, который расширяет базовый подход 3D-сплаттинга за счёт введения поля деформации. Поле деформации предназначено для описания временной эволюции каждой гауссианы, включая изменение позиции, ориентации, масштаба, цвета и непрозрачности. Этапы работы метода представлены на рисунке 1.

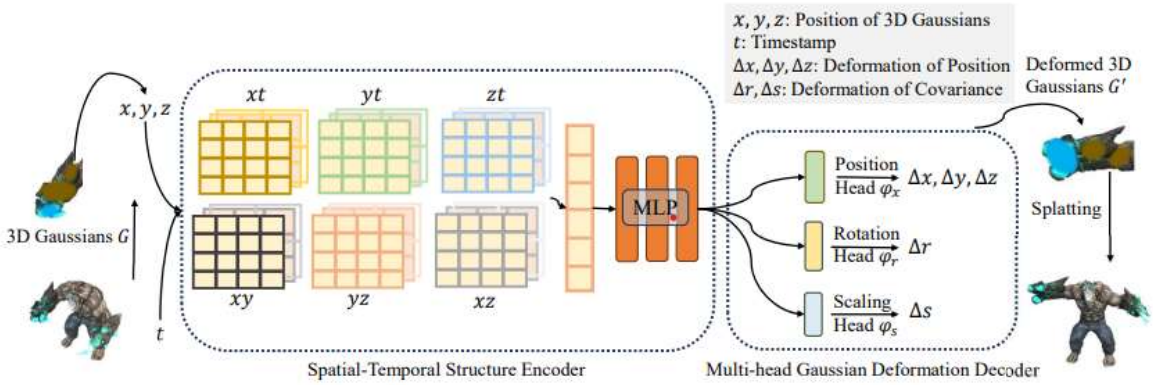


Рис. 1. Последовательность этапов работы модели

Поле деформации реализовано с помощью пространственно-временного энкодера и многоуровневого декодера деформации. Пространственно-временной энкодер H извлекает признаки из текущего состояния гауссиан G и временной метки t. Признаки извлекаются энкодером H, который учитывает как пространственное расположение, так и временную динамику гауссиан: $f_d = H(G,t)$.

Многоуровневый декодер деформации принимает на вход признаки и предсказывает изменения параметров гауссиан: $\Delta G = D(f_d) = \{\Delta X, \Delta r, \Delta s, \Delta C, \Delta\alpha\}$, где $\Delta X = \phi_x(f_d)$ — изменение позиции, $\Delta r = \phi_r(f_d)$ — изменение ориентации, $\Delta s = \phi_s(f_d)$ — изменение масштаба, $\Delta C = \phi_C(f_d)$ — изменение цвета, $\Delta\alpha = \phi_\alpha(f_d)$ — изменение прозрачности.

Итоговые параметры деформированной гауссианы выглядят так:
$(X', r', s', C', \alpha') = (X + \Delta X, r + \Delta r, s + \Delta s, C + \Delta C, \alpha + \Delta\alpha)$.

Применение 4D-сплаттинга

Для оценки эффективности рассмотренного метода проведены эксперименты на двух типах наборов данных: синтетическом «bouncingballs» и реальном «mipt-fathers». Синтетический набор данных представляет собой последовательность изображений, сгенерированных на основе трёхмерной компьютерной модели шаров. В данном случае положения камер детерминированы и доступны в метаданных датасета [7]. Реальные данные получены путём видеосъёмки с мобильной камеры. Записанное видео было разбито на отдельные кадры. Позиции камеры в трёхмерном пространстве для каждого кадра, а также облако точек трёхмерной сцены [6], были восстановлены с применением метода structure-from-motion COLMAP [8].

Оценка экспериментальных результатов проводилась с использованием следующих метрик: пиковое отношения сигнала к шуму PSNR, индекс структурного сходства SSIM [9], многоуровневый индекс структурного сходства MS-SSIM, индекс структурного различия D-SSIM, метрика схожести фрагментов изображений, основанная на перцептивном обучении, LPIPS [10].

Результаты приведены в таблице 1.

Таблица 1

Значения метрик

№	bouncing balls	mipt-fathers
PSNR	40.85	30.51

SSIM	0.9944	0.9202
MS-SSIM	0.9956	0.9810
D-SSIM	0.0022	0.0095
LPIPS	0.0057	0.0539
FPS (рендеринг видео)	—	~52
FPS (рендеринг 160 кадров)	171.8	—
FPS (рендеринг 20 кадров)	77.7	—
Количество гауссиан	27382	359 908

Значения метрик, полученные для синтетического датасета, свидетельствуют о высоком уровне качества реконструкции. Для реального датасета значения метрик ниже, однако остаются на достаточно высоком уровне, что подтверждает хорошее качество реконструкции, несмотря на сложность реальной сцены и наличие артефактов, характерных для подобного типа данных.

Общее время обучения модели на синтетическом датасете составило 4 минуты 52 секунды, а на реальном — 33 минуты 27 секунд. Следует отметить, что реальная сцена значительно превосходит синтетическую по уровню детализации. Это приводит к большему количеству гауссиан, необходимых для представления сцены, и существенной разнице во времени обучения нейронной сети, вычисляющей поле деформации.

Производительность 4D-сплаттинга была протестирована на различных сценариях рендеринга. Скорость рендера 20 кадров составила 77.7 FPS для синтетического датасета, при увеличении числа кадров до 160 скорость возросла до 171.8 FPS, что связано с параллелизацией на уровне GPU. Во время рендера полной видеопоследовательности реальной сцены из 272 кадров достигнута скорость ~52 FPS, что позволяет использовать метод в режиме реального времени.

Список литературы

1. Kerbl B. et al. 3d gaussian splatting for real-time radiance field rendering //ACM Trans. Graph. – 2023. – Т. 42. – №. 4. – С. 139:1-139:14.
2. Chen G., Wang W. A survey on 3d gaussian splatting //arXiv preprint arXiv:2401.03890. – 2024.
3. Zwicker M. et al. EWA volume splatting //Proceedings Visualization, 2001. VIS'01. – IEEE, 2001. – С. 29-538.
4. Wu G. et al. 4d gaussian splatting for real-time dynamic scene rendering //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2024. – С. 20310-20320.
5. Fu Y. et al. Colmap-free 3d gaussian splatting //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2024. – С. 20796-20805.
6. Schonberger J. L., Frahm J. M. Structure-from-motion revisited //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 4104-4113.
7. D-NeRF: Neural Radiance Fields for Dynamic Scenes // GitHub.com [Электронный ресурс]. URL: https://github.com/albertpumarola/D-NeRF (дата обращения: 28.04.2025).
8. D-NeRF: COLMAP // GitHub.com [Электронный ресурс]. URL: https://github.com/colmap/colmapNeRF (дата обращения: 28.04.2025).
9. Wang Z. et al. Image quality assessment: from error visibility to structural similarity //IEEE transactions on image processing. – 2004. – Т. 13. – №. 4. – С. 600-612.
10. Zhang R. et al. The unreasonable effectiveness of deep features as a perceptual metric //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – С. 586-595.

11. Mildenhall B. et al. Nerf: Representing scenes as neural radiance fields for view synthesis //Communications of the ACM. – 2021. – Т. 65. – №. 1. – С. 99-106.

АВТОМАТИЗИРОВАННЫЙ РАСЧЕТ ПОКАЗАТЕЛЕЙ САГИТТАЛЬНОГО БАЛАНСА ПО ТРЕХМЕРНОЙ МОДЕЛИ ПОЗВОНОЧНИКА, ПОСТРОЕННОЙ ИЗ РЕНТГЕНОГРАММ С ПОМОЩЬЮ АЛГОРИТМОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Угольникова Е.А.[1,2]

[1]Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»,
[2]Институт вычислительной математики и математической геофизики СО РАН (ИВМиМГ СО РАН)

Аннотация. В работе представлены результаты разработки автоматизированного расчета показателей сагиттального баланса по трехмерной визуализации позвоночника, построенной в ПО 3D Slicer. Представленная трехмерная визуализация позвоночника построена по двум рентгеновским изображениям в боковой и фронтальной проекции с применением моделей глубокого обучения YOLO и U-Net. Рассчитывались основные шесть показателей сагиттального баланса (SVA, SS, PT, PI, TK, GLL), необходимые для первичного анализа позвоночника в области вертебрологии, травматологии и ортопедии.

Ключевые слова: машинное обучение, обработка медицинских изображений, рентгенограмма, сагиттальный баланс

Современная медицина научилась лечить многие смертельные болезни, но теперь перед ней стоит новая важная задача — помогать людям жить полноценно, несмотря на хронические заболевания, которые не угрожают жизни, но серьезно ухудшают её качество. Например, болезни позвоночника — такие как остеохондроз, сколиоз или межпозвонковые грыжи — часто вызывают постоянную боль, ограничивают движение и затрудняют даже простые действия. Ключевым элементом их диагностики и лечения является сагиттальный баланс — оптимальное выравнивание позвоночника в вертикальной плоскости, нарушение которого провоцирует боль, дегенеративные изменения и ограничение подвижности. Однако ручной расчет десятков параметров баланса (таких как SVA, PI, PT) остается трудоемким и субъективным, что ограничивает их применение в клинической практике. Автоматизация этих расчетов на основе методов ИИ и компьютерного зрения становится критически важной для повышения точности диагностики, персонализации лечения и, как следствие, улучшения качества жизни пациентов.

Более 40 параметров сагиттального баланса делятся на две группы. К позвоночно-тазовым индексам относятся PI (тазовый индекс), PT (наклон таза), SS (наклон крестца), GLL (поясничный лордоз) и TK (грудной кифоз). Параметры глобального баланса включают SVA (сагиттальная ось), SSA (позвоночно-крестцовый угол) и др. Дополнительные показатели, такие как SFD, C2-C7 offset, LL и другие, детализированы в [1], обеспечивают комплексную оценку биомеханики позвоночника. Несмотря на большое количество этих параметров, во врачебной практике в основном используются около 10 основных, а расчет всех сорока занимает чрезмерное количество времени. Поэтому автоматизированный расчет этих параметров существенно упростит и ускорит работу врачей рентгенологов и вертебрологов, а также расширит границы исследований.

В представленной работе приведен автоматизированный расчет основных показателей сагиттального баланса. Расчет показателей проводится по формулам, описанным в [1].

Для расчета показателей используется разработанный дополнительный модуль для ПО 3D Slicer, в котором с использованием моделей глубокого обучения YOLO [2] для детекции позвонков на фронтальной и боковой рентгенограмме и U-Net [3] для сегментации позвонков на детектированных областях. По полученным результатам строится трехмерная визуализация позвоночника, на основании которой и происходит расчет показателей сагиттального баланса. После построения трехмерной визуализации от пользователя требуется отметить на входных рентгенограммах по три точки на каждой головке бедренных костей, по трем точкам находится центр головки бедренной кости, а затем находится середина отрезка, соединяющего две головки бедренных костей в боковой проекции. Эта точна является одним из важнейших ориентиров для расчетов, не предусмотренная встроенными в модуль алгоритмами.

На рисунке 1 показан пример входных данных для модуля.

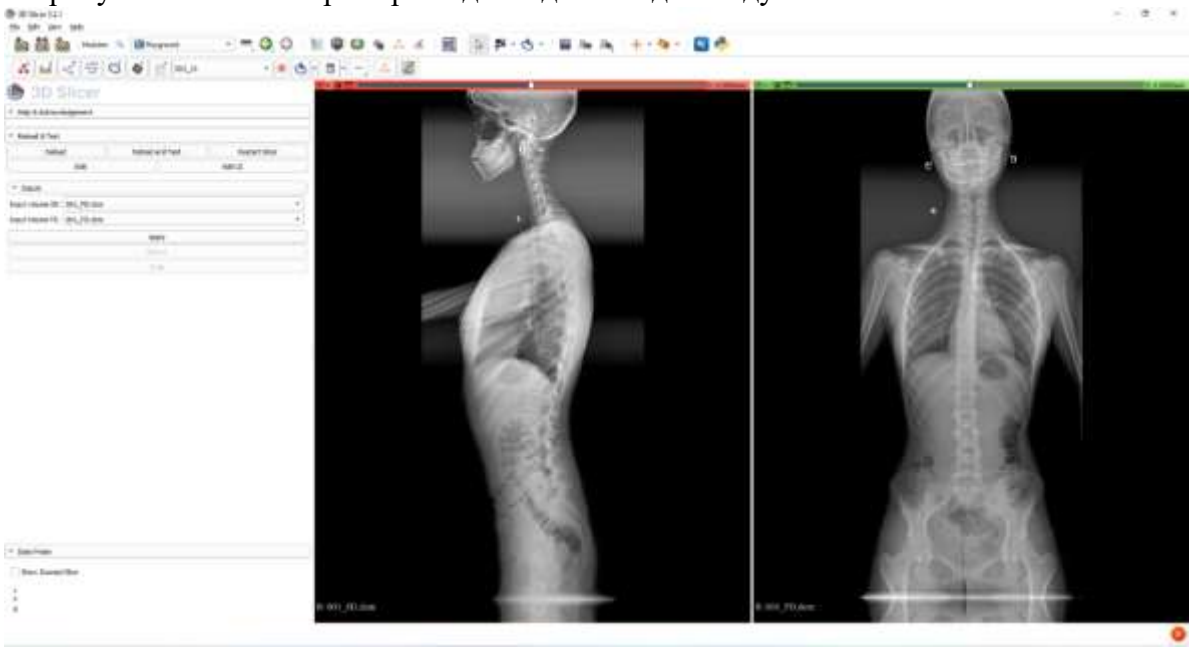


Рис. 1. Интерфейс и пример входных данных модуля 3D Slicer

На рисунке 2 показан результат трехмерной визуализации позвоночника, состоящий из простых параллелепипедов. А на рисунке 3 приведен результат расчета и отображения показателей сагиттального баланса.

Разработанное решение для автоматизированного расчета параметров сагиттального баланса демонстрирует высокую клиническую значимость, объединив методы компьютерного зрения и глубокого обучения. Программный комплекс обеспечивает точное определение ключевых показателей (SVA, PI, PT, SS) на основе бипланарных рентгенограмм, сокращая время анализа с часов до минут. Интеграция алгоритма с платформой 3D Slicer позволяет визуализировать результаты в режиме реального времени. Представленное решение открывает новые возможности для персонализированной вертебрологии и улучшения качества жизни пациентов.

Работа выполнена в рамках государственного задания Института вычислительной математики и математической геофизики СО РАН (проект № 0251-2022-0002).

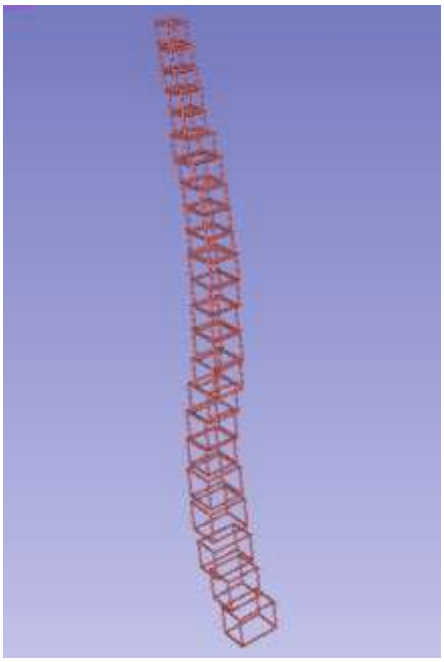


Рис. 2. Пример построенной 3D модели.

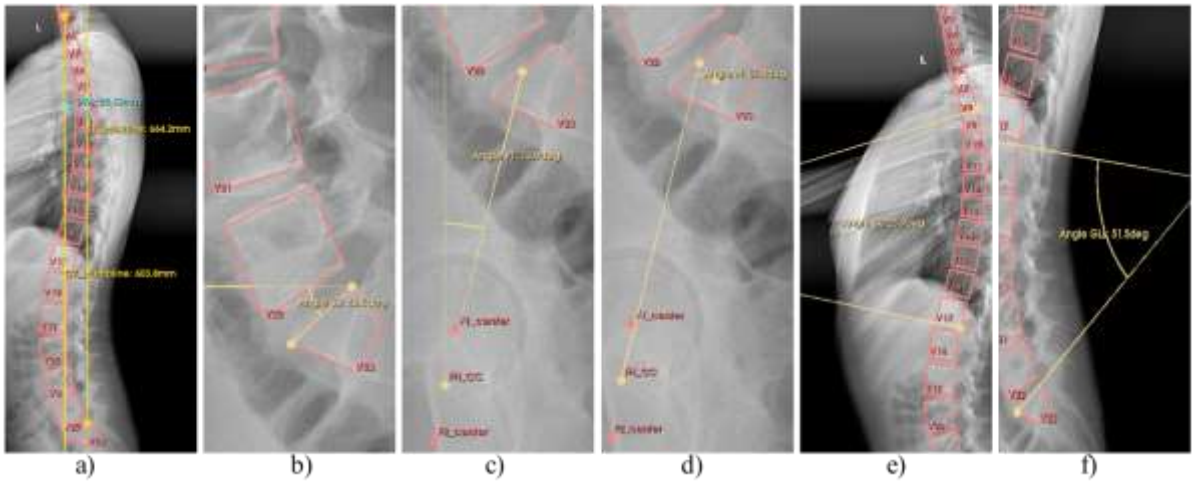


Рис. 3. Пример рассчитанных параметров. (Рассчитанные параметры: a) SVA (sagittal vertical axis, сагиттальная вертикальная ось), b) SS (sacral slope, наклон крестца), c) PT (pelvic tilt, наклон таза), d) PI (pelvic incidence, тазовый индекс), e) TK (thoracic kyphosis, грудной кифоз), f) GLL (global lumbar lordosis, общий поясничный лордоз))

Особая благодарность Новосибирскому научно-исследовательскому институту травматологии и ортопедии им. Я.Л. Цивьяна за разметку и предоставление данных для исследования.

Список литературы

1. А. В. Крутько, В. В. Рерих, В. М. Прохоренко и О. Н. Леонова, Нарушения сагиттального баланса при заболеваниях и повреждениях позвоночника: учебное пособие для врачей, Новосибирск: ИПЦ НГМУ, 2020, p. 80.
2. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," vol. 5, 2016.

3. O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," vol. 5, 2015.

АРХИТЕКТУРА СОВРЕМЕННЫХ СИСТЕМ НА ОСНОВЕ LLM

М.А. ХАРИТОНОВ

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В исследовании рассматриваются существующие архитектурные решения для LLM-систем и сравнительный анализ различных подходов их оптимизации. Основные критерии для оптимизации: качество ответов и время обработки запросов, масштабируемость и эффективность использования ресурсов. В результате были найдены оптимальные архитектуры систем. Результаты работы могут быть использованы в дальнейшем при разработке и интеграции решений на основе LLM.

Ключевые слова: система, архитектура, LLM, нейронные сети, дообучение, RAG, мониторинг, оркестровка, защитный UX, кэширование, LoRA, AI-агент.

Введение

Архитектура систем, построенных на основе Large Language Models (LLM) [12], должна сочетать в себе мощь глубокого обучения, масштабируемость и гибкость, что позволит интегрировать их в различные сферы, включая бизнес, образование, медицину и развлечения. Однако проектирование таких систем требует учета множества аспектов: от выбора подходящей модели и оптимизации производительности до обеспечения безопасности и этичности использования. В данной работе рассматриваются ключевые принципы архитектуры современных приложений на основе LLM, их компоненты, технические вызовы и возможные решения.

Анализ существующих систем

Рассмотрим примеры использования LLM в системах, выделяя основные архитектурные шаблоны.

1. Компания SL Soft [1] применила связку LLM и Retrieval-Augmented Generation (RAG) [13] для автоматизации первой линии поддержки клиентов, с помощью которой анализирует запросы пользователей, учитывая контекст диалога и намерения. Доступ LLM посредством RAG к актуальным корпоративным данным позволил сократить время обработки запросов на 50–70%, снизить нагрузку на операторов и минимизировать число ошибок. RAG предоставляет доступ к данным из БД через векторную БД. Взаимодействие векторной БД с LLM происходит через промежуточный слой бизнес-правил.

Для оркестровки данных вместо устоявшегося Airflow с более развитой инфраструктурой был выбран Dagster. Плюсами такого выбора стали: простота, гибкость работы с различными конвейерами и продвинутая визуализация данных.

LLM выбиралась под специфику задачи исходя из точности классификации, размера и стоимости из списка наиболее мощных, однако итоговая модель не была протестирована для других задач. Эталонным результатом модели обычно является проверка на MMLU (Massive Multitask Language Understanding) по различным областям знаний.

2. Чат-бот компании DoorDash поддерживает высокое качество ответов благодаря системе ограждений LLM Guardrail [2]. Она предотвращает “галлюцинации” модели и

фильтрует ответы, противоречащие внутренним нормам. Если ответ модели неудовлетворителен, то происходит обратная связь от эксперта (метод RLHF [3]).

Для последующей оценки уровня работы чат-бота используется LLM Judge, который анализирует качество ответов по пяти метрикам: корректность поиска информации, точность ответа, грамматическая правильность и стилистика, логическая согласованность и релевантность запросу.

3. Проект presentsimple.ai [4] разработал сервис для автоматического создания презентаций на основе текстовых запросов. Данный сервис использует AI-агентов – это автономные системы на основе языковой модели или ансамбля моделей, которые выполняют свою специфическую задачу по техническому заданию (ТЗ).

При ограниченных вычислительных ресурсах для полного дообучения и для большей гибкости системы были выбраны методы PEFT (Parameter-Efficient Fine-Tuning) [5]. Для адаптации LLM под различные задачи использовались LoRA-адаптеры (Low-rank adaptation) [6]. По сравнению с полным переобучением они быстрее обучаются, снижают затраты на вычисление и хранение, их удобно развёртывать, а ещё позволяют избегать катастрофического забывания [7]. Это обеспечило снижение вычислительных ресурсов при сохранении высокого качества генерируемых презентаций. Есть и другие методы PEFT, однако в большинстве задач LoRA их превосходит по эффективности [10].

Важно отметить, что есть возможность использовать гибридный подход с использованием RAG и дообучения. Однако, presentsimple.ai не использует RAG и обходится одним дообучением, так как почти нет необходимости в обращении к внешним знаниям за счёт наличия достаточного объёма тренировочных данных и из-за высокой специализации:

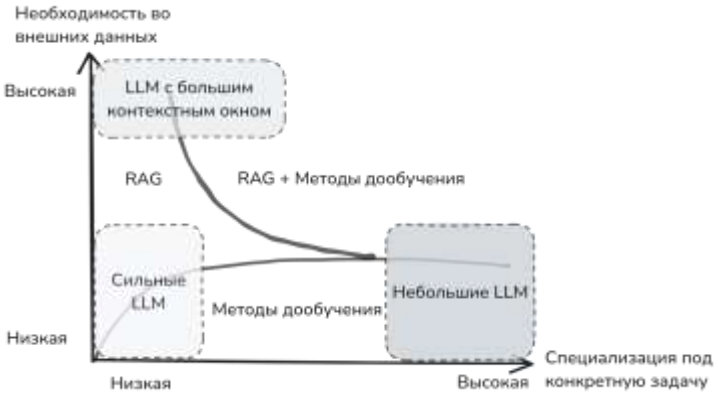


Рис. 1. Выбор между RAG и методами дообучения

На рис. 1 также отражена связь методов с оптимальным выбором LLM. Например, “сильные” LLM, которые обладают достаточно высокими показателями (MMLU и др.), будут не так нуждаться в доступных данных в целом, однако метод RAG может быть чуть более эффективным за счёт более уверенных знаний модели о мире. LLM с большим контекстным окном позволяют держать в памяти одновременно больше актуальной информации об интересующей области. Небольшие же LLM больше подвержены плохим ответам, что делает полноценное дообучение более актуальным.

Помимо изложенных идей, стоит учитывать, что RAG лучше подходит для минимизации галлюцинаций, при динамических данных и для интерпретируемости.

Возможно комбинировать несколько методов, однако тогда преимущество в скорости и уменьшении затрат может быть не таким явным.

QLoRA (LoRA с квантованием) [8] показывает схожую с LoRA производительность, имеет примерно на 75% меньший пик использования GPU, но на 66% медленнее в обучении и из-за накладных расходов на квантование и до 40% дороже по эффективности использования ресурсов [9]. DoRA [11] является развитием LoRA, однако из-за относительной новизны примеров её использования для LLM недостаточно.

Критерии оптимальной архитектуры

Исходя из рассмотренных примеров, можно выделить набор основных критериев оптимальной архитектуры систем на основе LLM, среди которых: небольшое время обработки запросов, качество ответов (удовлетворение бизнес-правилам), многозадачность LLM, мониторинг качества LLM, эффективная интеграция данных, эффективное дообучение и использование обратной связи.

Согласно предложенным критериям, можно рассмотреть архитектурные шаблоны, которые могут быть адаптированы и под другие проекты. Расположим эти шаблоны на осях основных компонентов. В любой архитектуре на основе LLM можно выделить 4 основные компоненты: близость к пользователю, близость к данным, повышение эффективности/функциональности или уменьшение расходов/рисков. Расположим основные шаблоны на карте позиционирования:

Рис. 2. Распределение шаблонов на осях основных компонент.

Кроме рассмотренных шаблонов, не менее важными, ориентированными на комфорт пользователя компонентами являются защитный UX и кэширование:

- Защитный UX помогает избежать пользователям нежелательных последствий при взаимодействии с LLM, включая разъяснение границ модели, возможность изменения запроса или контроль качества ввода, тем самым улучшая пользовательский опыт [12].

- Кэширование позволяет использовать уже вычисленные результаты (например, ответы модели или данные из базы знаний для RAG), а за счёт быстроты кэша достигается большая производительность по сравнению с традиционными БД (PostgreSQL и др.), тем самым уменьшается нагрузка на всю систему и повышается её масштабируемость и скорость ответов.

Формирование рекомендаций

В зависимости от ресурсов/целей организации возможно частичное использование представленных на рис. 2 шаблонов.

Организациям, стремящимся как можно быстрее добиться наибольшей эффективности и протестировать технологии, имеет смысл сфокусироваться на правой части карты позиционирования, а для привлечения пользователей – на левой.

Если организация обладает достаточными ресурсами, то а для полноценного развертывания системы имеет смысл рассматривать доработку системы шаблонами из левой части карты.

Другие рекомендации на основе проведённого анализа:

1. Использовать оркестровку данных с RAG для актуальности информации и быстрой обработки запросов.
2. Выбирать модели с учетом оценки на MMLU, объема данных и доступных ресурсов, а тестировать под конкретные задачи.
3. Применять PEFT-методы (LoRA и др.) для экономии вычислительных мощностей.
4. Внедрять обратную связь (RLHF) и мониторинг для улучшения качества ответов.
5. Использовать ограждения (Guardrails) и защитный UX для повышения безопасности и надежности модели.
6. Применять кэширование и квантование (QLoRA) для оптимизации затрат.
7. Комбинировать RAG и дообучение для баланса актуальности данных и настройку под задачи, но перед этим желательно проверить требования к системе.

Заключение

В ходе исследования были найдены и формализованы оптимальные архитектурные решения для различных требований к системе. Дальнейшим шагом исследователя является проверка архитектур на практике и их дальнейшая оптимизация.

Список литературы

1. Крюков М. Применение LLM + RAG для диалоговых систем в службе поддержки // Хабр. – 2025. – URL: https://habr.com/ru/companies/slsoft/articles/877914/ (дата обращения: 14.03.2025).
2. Малюгина Д. 10 примеров и вариантов использования RAG от реальных компаний. Со схемами и пояснениями // Хабр. – 2025. – URL: https://habr.com/ru/articles/886732/ (дата обращения: 14.03.2025).
3. Ouyang L. et al. Training language models to follow instructions with human feedback //Advances in neural information processing systems. – 2022. – Т. 35. – С. 27730-27744.
4. Фетисов П. LoRA fine-tuning для генерации презентаций: как мы выбирали метод дообучения LLM в presentsimple.ai // Хабр. – URL: https://habr.com/ru/companies/yandex_cloud_and_infra/articles/884336/ (дата обращения: 14.03.2025).
5. Ding N. et al. Parameter-efficient fine-tuning of large-scale pre-trained language models //Nature Machine Intelligence. – 2023. – Т. 5. – №. 3. – С. 220-235.
6. Hu E. J. et al. Lora: Low-rank adaptation of large language models //ICLR. – 2022. – Т. 1. – №. 2. – С. 3.
7. Батурова Д. Parameter-Efficient Fine-Tuning (PEFT): методы LoRA, Prefix tuning, Prompt tuning и Adapters // Хабр. – URL: https://habr.com/ru/articles/791966/ (дата обращения: 14.03.2025).
8. Dettmers T. et al. Qlora: Efficient finetuning of quantized llms //Advances in neural information processing systems. – 2023. – Т. 36. – С. 10088-10115.
9. LoRA и QLoRA: эффективная адаптация больших языковых моделей // Google Cloud. – URL: https://cloud.google.com/vertex-ai/generative-ai/docs/model-garden/lora-qlora (дата обращения: 14.03.2025).
10. Беляков К. Оптимизируем дообучение LLM: теория + гайд // Хабр. – URL: https://habr.com/ru/articles/860892/ (дата обращения: 14.03.2025).
11. Liu S. Y. et al. Dora: Weight-decomposed low-rank adaptation //Forty-first International Conference on Machine Learning. – 2024.

12. Guidelines for Human-AI Interaction // Microsoft Research. – URL: https://www.microsoft.com/en-us/research/project/guidelines-for-human-ai-interaction/ (дата обращения: 14.03.2025).
12. Yao Y. et al. A survey on large language model (llm) security and privacy: The good, the bad, and the ugly //High-Confidence Computing. – 2024. – С. 100211.
13. Lewis P. et al. Retrieval-augmented generation for knowledge-intensive nlp tasks //Advances in neural information processing systems. – 2020. – Т. 33. – С. 9459-9474.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ВЫЖИВАЕМОСТИ ПРИ РАКЕ МОЛОЧНОЙ ЖЕЛЕЗЫ

Чернов Е.А.[1], Мелдо А.А.[1,2], Агаева А.В.[2], Богданов М.А.[2], Вальков М.Ю.[2]

[1]Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина»)
[2]ФГБОУ ВО "Северный государственный медицинский университет" МЗ РФ

Аннотация. В работе сравнивались три модели прогнозирования выживаемости пациенток с раком молочной железы: регрессия Кокса, случайный лес выживаемости и DeepHit. Использованы данные 8 959 пациентов из популяционного регистра Архангельской области за 2000–2022 годы с тремя типами исходов. Модели машинного обучения показали более высокую точность по сравнению с классической моделью. Также выделены ключевые факторы, влияющие на выживаемость.

Ключевые слова: рак молочной железы, прогнозирование выживаемости, конкурирующие риски, машинное обучение, DeepHit, RSF, Cox

Введение

Прогнозирование выживаемости пациентов является одной из ключевых задач в медицинских исследованиях, особенно в онкологии, где точность прогноза напрямую влияет на выбор тактики лечения. Под выживаемостью понимают долю пациентов, остающихся в живых к определенному моменту времени от начала наблюдения, а основным объектом анализа выступает время до наступления события (например, смерти или рецидива) [1]. Традиционно для анализа выживаемости применяются статистические методы, наиболее распространена модель пропорциональных рисков Кокса (Cox PH) [2]. Однако ее использование предполагает соблюдение ряда допущений, что не всегда выполняется в реальных клинических условиях. Методы машинного обучения [3], такие как Random Survival Forests (RSF) [4] и DeepHit [5], демонстрируют более высокую точность за счёт способности учитывать сложные нелинейные зависимости и взаимодействия признаков.

Также принципиально важной задачей является определение ключевых факторов риска, влияющих на исход заболевания. Особое значение это приобретает при анализе структуры смертности онкологических пациентов, где существенную роль играют сопутствующие причины летальности. Как показывают авторы исследования [6], проблема предотвращения неонкологической смертности у больных раком требует особого внимания поскольку на показатели выживаемости могут существенно влиять особенности пациента (сопутствующие заболевания, осложнения, и т. д.). Наиболее полные данные для решения таких комплексных задач содержатся в популяционных регистрах рака, которые объединяют детализированную клиническую информацию с длительным наблюдением [7].

Именно такие репрезентативные данные позволяют разрабатывать точные прогностические модели, способные дифференцировать риски различной природы.

В данном исследовании сравнивается эффективность трёх подходов: модели Кокса, RSF и DeepHit. Особое внимание уделяется комплексному решению двух взаимосвязанных задач — анализу выживаемости и определению значимых факторов риска. Анализ выполнен на репрезентативной выборке, включающей детализированные клинические данные 8 959 пациентов с раком молочной железы (РМЖ) из популяционного регистра Архангельской области, что обеспечивает высокую клиническую достоверность полученных результатов.

Характеристика данных и методы анализа

Анализ основан на 42 клинико-демографических признаках (включая исходы заболевания) пациентов Архангельского онкологического диспансера за 2000–2022 гг. Все данные были анонимизированы. В качестве конкурирующих исходов рассматривались три причины смерти: смерть от РМЖ (n=2729), от других онкозаболеваний (ДОЗ) (n=261) и от сердечно-сосудистых заболеваний (ССЗ) (n=766); остальные 5203 случая — цензурированные. Данные прошли проверку и предварительную обработку, включая кодирование категориальных признаков и стратифицированное разделение на обучающую и тестовую выборки.

В моделях Cox PH и RSF для каждой причины смерти строилась отдельная модель с трактовкой остальных причин как цензурированные наблюдения. DeepHit позволил одновременно моделировать все причины, и оценивать вероятность наступления событий используя кумулятивных функций заболеваемости (CIF), что обеспечивает согласованность прогнозов.

Сравнение моделей с и без учёта конкурирующих рисков допустимо при оценке их способности стратифицировать пациентов по уровню риска. Однако при интерпретации значимости признаков следует учитывать методологические различия: модели без учёта конкуренции оценивают влияние фактора на интересующий исход изолированно, тогда как модели с учётом конкурирующих событий отражают его вклад с поправкой на вероятность альтернативных исходов.

Для оценки устойчивости моделей применялись кросс-валидация и бутстрэппинг. Алгоритмы RSF и DeepHit оптимизировались с помощью Optuna, а важность признаков оценивалась методом пермутаций. Сравнение проводилось по C-индексу, интегральной ошибке Брайера (IBS) и AUC на горизонтах 3, 7 и 12 лет, отражающих ключевые этапы динамики выживаемости. Во всех таблицах квадратные скобки обозначают 95% доверительные интервалы [нижний; верхний] для соответствующих показателей.

Результаты модели Cox Proportional-Hazards

Перед обучением была проведена стратификация признаков, для которых допущение пропорциональности рисков не подтвердилось. Это позволило сохранить прогностическую способность модели, но ограничило интерпретацию вклада этих переменных.

Cox-модель продемонстрировала ограниченные прогностические возможности. Для модели с причиной смерти от РМЖ значение C-index составило 0.59, что свидетельствует о слабой способности различать пациентов по риску. Наиболее значимыми факторами стали размер опухоли (T по TNM) и распространение за пределы одной локализации. Для других исходов — смерть от ДОЗ (C-index = 0.63) и смерть от ССЗ (C-index = 0.57) — дискриминационная способность моделей также оказалась низкой, а статистически значимых

факторов выявлено не было. Следует отметить, что расчёт дополнительных метрик прогностической точности (IBS, AUC) был затруднён в связи с тем, что используемая библиотека не поддерживает оценку таких метрик для стратифицированных моделей.

Результаты модели Random Survival Forest

RSF показал высокую прогностическую точность для причин РМЖ и ССЗ (C-index > 0.82, AUC > 0.80) с устойчивыми оценками и низкими значениями IBS. Ведущими факторами риска смерти от РМЖ стали компоненты TNM-классификации и возраст. Для риска смерти от ДОЗ точность была ниже (C-index 0.65), а предикторы исходов, связанных с ССЗ и ДОЗ в целом совпадали: ключевым фактором оставался возраст, тогда как тип хирургического лечения, менопаузальный статус, поражение лимфоузлов и задержка терапии играли второстепенную роль.

Таблица 1

Причина смерти	C-index	IBS	AUC (3 года)	AUC (7 лет)	AUC (12 лет)
РМЖ	0.844, [0.827-0.859]	0.140, [0.128-0.152]	0.885, [0.862–0.906]	0.860, [0.836–0.882]	0.818, [0.787–0.844]
ДОЗ	0.646, [0.553-0.730]	0.043, [0.033-0.056]	0.725, [0.574–0.848]	0.651, [0.533–0.772]	0.680, [0.586–0.776]
ССЗ	0.825, [0.785-0.860]	0.076, [0.065-0.088]	0.848, [0.782–0.901]	0.842, [0.799–0.883]	0.861, [0.813–0.902]

Результаты модели DeepHit

Модель DeepHit продемонстрировала наилучшие результаты: для риска смерти от РМЖ и ДОЗ значения C-index и AUC оказались немного выше, чем у RSF, а от ССЗ — несколько ниже. При этом более узкие доверительные интервалы указывают на высокую точность модели. Для летальность от РМЖ основное влияние оказывали стадия, TNM-T, возраст и задержка начала лечения. В прогнозе летальности от ССЗ доминировал возраст, тогда комбинированные методы лечения и стадия заболевания играли меньшую роль. Несмотря на умеренный C-index (0.63), модель обеспечила лучшие метрики и для исходов, связанных с ДОЗ, где наибольшее значение имели возраст, локализация опухоли в верхне-наружном квадранте и отсутствие лечения оказывали меньшее влияние.

Таблица 2

Причина смерти	C-index	IBS	AUC (3 года)	AUC (7 лет)	AUC (12 лет)
РМЖ	0.850, [0.842-0.857]	0.131, [0.126-0.136]	0.892, [0.882-0.903]	0.860, [0.850-0.871]	0.851, [0.837-0.866]
ДОЗ	0.629, [0.592-0.667]	0.052, [0.045-0.058]	0.739, [0.674-0.799]	0.719, [0.676-0.761]	0.776, [0.736-0.815]
ССЗ	0.807, [0.787-0.826]	0.090, [0.083-0.097]	0.836, [0.803-0.864]	0.837, [0.815-0.859]	0.849, [0.828-0.869]

Выводы

Сравнительный анализ показал, что модель на основе DeepHit демонстрирует наилучшие результаты и может считаться наиболее перспективной для анализа нелинейных данных с конкурирующими рисками, что делает её ценной для использования в реальной клинической практике.

Для анализа факторов риска применялись модели, отражающие как изолированное, так и комбинированное влияние признаков. Такой анализ позволяет более полно оценить структуру факторов риска и является важным условием для построения надёжных прогностических моделей. Сходимость результатов подтверждает устойчивость выявленных закономерностей. Наиболее значимыми факторами оказались компоненты TNM-классификации, возраст и задержка лечения, что соответствует современным клиническим наблюдениям.

Полученные результаты подтверждают, что современные методы машинного обучения способны существенно повысить точность прогноза выживаемости и выявлять значимые клинические закономерности, открывая путь к более обоснованным и эффективным решениям в онкологической практике.

Список литературы

1. Беляев А. М., Михнин А. Е., Рогачев М. В. Подготовка данных и анализ выживаемости в пакетах статистических программ MedCalc и Statistica: учебное пособие для обучающихся в системе высшего и дополнительного профессионального образования. – Санкт-Петербург: НМИЦ онкологии им. Н.Н. Петрова, 2022. – 56 с.
2. Cox D. Regression models and life-tables // Journal of the Royal Statistical Society. Series B (Methodological). – 1972. – Vol. 34. – P. 187–220.
3. Wang P., Li Y., Reddy C. Machine learning for survival analysis: A survey // ACM Computing Surveys. – 2019. – Vol. 51. – P. 1–36.
4. Ishwaran H., Kogalur U.B., Blackstone E.H., Lauer M.S. Random survival forests // The Annals of Applied Statistics. – 2008. – Vol. 2, No. 3. – P. 841–860. – DOI: 10.1214/08-AOAS169.
5. Lee C., Zame W., Yoon J., van der Schaar M. DeepHit: A Deep Learning Approach to Survival Analysis With Competing Risks // Proceedings of the AAAI Conference on Artificial Intelligence. – 2018. – Vol. 32, No. 1. – DOI: 10.1609/aaai.v32i1.11842.
6. Вальков М. Ю., Гржибовский А. М., Кудрявцев А. В., Хафизов Ф. Р., Петров И. Н. Использование искусственного интеллекта для прогнозирования и предотвращения неонкологической смертности у онкологических больных: протокол исследования АРИЛИС // Экология человека. – 2024. – Т. 31, № 4.
7. Вальков М. Ю., Карпунов А. А., Коулман М. П., Голубева Н. М., Гудкова М. В., Давыдов М. И., Джанджглава М. Г., Жуков Н. В., Иванова Г. И., Киселёв Ю. М., Ларина В. И., Мартынов С. Н., Петров И. Н., Поддубная И. В., Прокопчук А. Ю., Сидоренко В. И., Тюрин И. А., Хафизов Ф. Р., Шевченко О. С. Популяционный раковый регистр как ресурс для науки и практического здравоохранения // Экология человека. – 2017. – Т. 24, № 5. – С. 54–62.

РЕШЕНИЕ ЗАДАЧИ АВТОМАТИЧЕСКОГО ГЛОССИРОВАНИЯ ПРИ МИНИМАЛЬНОМ ОБЪЁМЕ ДАННЫХ: ПРИМЕР ИЖОРСКОГО ЯЗЫКА

ЧЕПАСОВ Д.В., ЩЕДРИН А.А.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Аннотация. В работе представлены результаты обучения нейросети глоссированию русского и ижорского языков. Изучено влияние объёма данных на возможности перевода и глоссирования текста на различных языках.

Ключевые слова: Перевод, Искусственный интеллект, Большие лингвистические модели, Ижорский язык, Объём обучающей выборки, глоссирование.

В современном мире проблема сохранения языкового разнообразия стоит особенно остро. Согласно данным «ЮНЕСКО», исчезновение одного языка происходит в среднем

каждые 14 дней [1]. Данная статистика подчёркивает необходимость поиска решения поддержки исчезающих языков. Машинное обучение — один из многообещающих инструментов в исследуемой области. Способность алгоритмов машинного обучения извлекать закономерности из ограниченных объемов данных открывает новые перспективы для лингвистических исследований. Правильно обученные нейронные сети могут стать ключом к решению проблемы даже в условиях дефицита информации.

В лингвистике при изучении малых языков прибегают к методу глоссирования. Глоссирование языка — процесс разъяснения или интерпретации лексических единиц, грамматических структур и синтаксических особенностей текста, чаще всего на другом языке или с привлечением научной терминологии. В научном контексте глоссирование представляет собой подробное аннотирование и лингвистический анализ текста, направленный на выявление его структуры, смысла и функций языковых элементов.

Глоссирование может преследовать разные цели в зависимости от дисциплины, однако в лингвистике глоссирование прежде всего используется для анализа текстов на малоизвестных, недостаточно изученных языках, в том числе для языков, находящихся под угрозой исчезновения. Одним из ключевых инструментов является **межлинейный перевод (interlinear gloss)** — формат представления текста, в котором каждая единица оригинала (слово или морфема) сопровождается пояснением лексического значения и/или грамматической функции.

Для однозначного представления структуры языка было принято решение придерживаться **Leipzig Glossing Rules** — международного признанного стандарта оформления глосс [2]. В соответствии с данным стандартом, лексические значения передаются строчными буквами (dog, run), а грамматические категории — прописными (PL, PAST, 1SG и т. д.), что обеспечивает однозначное и читаемое представление структуры языка, понятное как специалистам, так и не владеющими данным языком читателям. Используемый способ аннотирования позволяет не только интерпретировать содержание, но и выявлять типологические и морфосинтаксические особенности языка.

Большие языковые модели (LLM) демонстрируют высокую эффективность в различных задачах обработки естественного языка, включая машинный перевод, генерацию текста и синтаксический анализ. Важным преимуществом LLM является способность обрабатывать языки и диалекты, отсутствующие в обучающих данных, что открывает перспективы для документирования малых и исчезающих языков. Однако, несмотря на мощность, большинство LLM не оптимизированы под специфические задачи лингвистической аннотации, такие как автоматическое межлинейное глоссирование, где требуется точное соответствие морфем и грамматических значений.

При создании модели была взята за основу **GLOSSLM** [3], предложенная для автоматического межлинейного глоссирования (IGT) в условиях ограниченных данных. Модель основана на мультиязычной архитектуре ByT5, дообученной на крупнейшем существующем корпусе IGT-примеров, свыше 450 тыс. единиц. Модель демонстрирует устойчивую производительность даже для языков, не встречающихся в тренировочных данных, и особенно качественно справляется с глоссированием несегментированного текста, что критически важно в реальных задачах документирования. Однако в модели были выявлены существенные ограничения: несогласованность в обозначениях данных (одна и та же часть речи в разметке данных могла обозначаться разными способами), дисбаланс языкового обучения (большая часть модели была обучена на языке Arapaho).

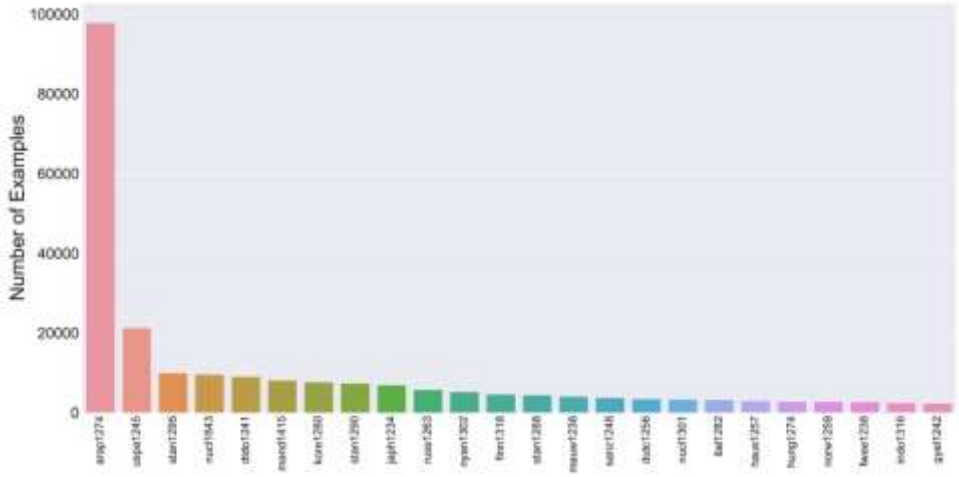


Рис. 1. Количество примеров на разных языках в работе Ginn (создано автором).

В работе использовалась предобученная модель google/T5, основанная на архитектуре Transformer и предназначенная для решения широкого круга задач NLP, включая машинный перевод. Был применен унифицированный text-to-text framework, где как входные, так и выходные данные представляются текстовыми строками. В отличие от моделей BERT-типа, ограниченных предсказанием классов или отрезков текста, данный подход обеспечивает большую гибкость за счет единой архитектуры для различных задач [4].

Исследование проводилось на базе модели GLOSSLM, доступной на платформе Hugging Face под идентификатором lecslab/glosslm. На первом этапе исследования модель была дообучена на русскоязычном корпусе, состоящем из 5762 пар предложение-глоссов. Результаты показали 90 % точность автоматического глоссирования русскоязычных текстов, что подтвердило эффективность выбранного подхода для языков с достаточным объемом обучающих данных. Для ижорского языка, относящегося к прибалтийско-финской группе уральской языковой семьи, близкородственному эстонскому и финскому языкам, был подготовлен корпус из 163 примеров, собранный в ходе полевых исследований с носителями языка. Точность модели была близка нулю, затем было принято решение использовать комбинированный подход: дополнительное применение BERT-модели для точечной подстановки значений. При обучении модели на ограниченном наборе данных точность предсказания составила около 10 %.

В ходе исследования была разработана система автоматического глоссирования на основе модели GLOSSLM (lecslab/glosslm). В ходе экспериментов были получены следующие результаты:

1) при обучении на русском корпусе с обучающей выборкой размером в 5762 примеров была достигнута 90 % точность глоссирования, из чего следует, что данного размера корпуса языка достаточно для обучения мультилингвальной модели;
2) при обучении ижорскому языку, используя выборку в 163 примера, показатели точности ненамного превысили уровень случайного угадывания, что свидетельствует о недостаточности данного объема данных для качественного обучения;
3) был сделан вывод, что на тысячах предложений-примеров языка возможно обучить модель, в то время как на десятках, как в случае с ижорским, почти невозможно.

Список литературы

1. Каждый месяц в мире умирает два языка: Российское образование [Электронный ресурс]. — 2013. — URL: https://www.edu.ru/news/education/24882/ (дата обращения 22.04.2025).
2. Comrie, Bernard. From the Leipzig Glossing Rules to the GE and RX lines. Corpus-based Studies of Lesser-described Languages: The CorpAfroAs corpus of spoken AfroAsiatic languages. — 2015. — PP. 207-219.
3. Michael Ginn, Lindia Tjuatja, Taiqi He, Enora Rice, Graham Neubig, Alexis Palmer, Lori Levin. GlossLM: A Massively Multilingual Corpus and Pretrained Model for Interlinear Glossed Text. EMNLP 2024 — 2024. — PP. 12267-12286.
4. Exploring Transfer Learning with T5: the Text-To-Text Transfer Transformer. Google Research. Adam Roberts, Colin Raffel. [Электронный ресурс]. — 2020. URL: https://research.google/ (дата обращения 22.04.2025).

СЕКЦИЯ АЛГОРИТМИЧЕСКАЯ МАТЕМАТИКА

МОДЕЛИРОВАНИЕ КОРОНАЛЬНЫХ ВЫБРОСОВ МАССЫ

АРУТЮНЯН С. Н.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В работе описывается реализация конусной модели корональных выбросов массы (КВМ) на основе наблюдений космической обсерватории SOHO/LASCO. Разработанная конусная модель используется в полуавтоматическом режиме и позволяет оценить кинетические и геометрические свойства произвольного КВМ. Также в работе описывается инструмент перевода полученных свойств КВМ в граничные условия для последующего МГД-моделирования. Результаты разработанной конусной модели с некоторым смещением повторяют стремления к пиковым значениям, замеченные при реальных измерениях КВМ, а также визуально и при более подробном сравнении схожи с результатами моделирования при использовании закрытой реализации конусной модели, предоставляемой агентством NOAA.

Ключевые слова: солнечный ветер, корональные выбросы массы, МГД-моделирование, конусная модель, прогноз космической погоды

Введение

Солнечным ветром называют поток плазмы, излучаемый из солнечной короны. Моделирование солнечного ветра является важной задачей, т.к. он может провоцировать сбои радиосвязи и навигации, вызывать магнитные бури на Земле, а также представлять опасность для космонавтов и космических аппаратов в открытом космосе. Моделируя поведение солнечного ветра на несколько суток вперед, можно спрогнозировать возможные опасности и принять меры для их предотвращения. В США прогнозированием занимается Центр прогноза космической погоды в составе Национального управления океанических и атмосферных исследований (Space Weather Prediction Center, NOAA SWPC) [1]. В России подобная служба была разработана в 2023 году [2]. В разработанной службе, подобно SWPC, для моделирования солнечного ветра и корональных выбросов массы используется МГД-модель [3].

Служба прогноза солнечного ветра

Используя уравнения магнитной гидродинамики (МГД), можно моделировать движение солнечной плазмы подобно движению жидкости. SWPC работает на основе такой трехмерной МГД-модели под названием ENLIL [3]. На момент начала работы над моделированием корональных выбросов массы (КВМ) уже была разработана программа, реализующая функциональность ENLIL. Она включает в себя библиотеку PLUTO [4], которая предоставляет численные методы решения системы уравнений МГД-уравнений. В работе [5] разработанная служба прогноза солнечного ветра была дополнена моделированием КВМ на основе кинематических и геометрических свойств КВМ, предоставляемых в открытом доступе организацией NOAA. В настоящей работе была реализована модель, позволяющая независимо оценивать кинематические и геометрические свойства КВМ с последующей интеграцией этих данных в граничные условия МГД-уравнений.

Конусная модель КВМ

В данной работе был сделан упор на изучение КВМ, используя изображения коронографов только с КА SOHO [6]. Для определения кинетических и геометрических свойств КВМ была выбрана аналитическая конусная модель (АКМ) [7]. АКМ базируется

на предположениях, что а) КВМ движется вплоть до солнечной короны с постоянным угловым размером, б) КВМ движется радиально и изотропно и в) источник КВМ находится вблизи солнечного экватора. Эти предположения имеют место только в пределах границы в 0.1 а.е. ($21.5R_S$). После этой границы в силу вступают предположения в рамках МГД-модели.

Конусная модель позволяет оценить кинетические и геометрические свойства КВМ по его двумерной проекции, видимой с коронографа LASCO (plane-of-sky). Для аппроксимации КВМ в данной работе моделируется трехмерный конус, основание которого совпадает с передним краем КВМ. Такой конус на plane-of-sky выглядит как эллипс с различными параметрами на изображении. Поэтому модель способна выполнить обратную задачу, т. е. по вручную отмеченному эллипсу смоделировать трехмерный конус, геометрические параметры которого совпадают с геометрическими параметрами наблюдаемого КВМ. Для определения скорости КВМ используется серия последовательных изображений с коронографа LASCO.

Численное моделирование с использованием результатов конусной модели

Для интеграции результатов конусной модели в граничные условия был разработан модуль **cone2bc**. В качестве начальных условий используются данные, полученные при наблюдении спокойного солнечного ветра.

В рамках cone2bc моделируется пересечение КВМ внутренней границы, находящейся на расстоянии $21.5R_s$ от центра солнца (R_s – солнечный радиус). Если предположить, что КВМ движется с постоянной скоростью и постоянным угловым размером, его пересечение с внутренней границей моделируется следующим неравенством:

$$(\theta - \theta_{CME})^2 + (\phi - \phi_{CME}) < \alpha(t),$$
$$\alpha(t) = \frac{\omega_{CME}}{2} sin[\frac{\pi}{2}\frac{t}{t_{half}}],$$
$$t_{half} = 21.5R_s \frac{\omega_{CME}}{2V_{CME}},$$

где θ_{CME}, ϕ_{CME} — широта и долгота направления распространения центра КВМ, ω_{CME} — угловой размер КВМ, V_{CME} — скорость КВМ.

Тогда для каждой точки, угловое расстояние до которой от точки $(\theta_{CME}, \phi_{CME})$ меньше значения $\alpha(t)$, в граничных условиях производятся следующие замены:

1. Плотность в точке становится равна среднему значению плотности солнечного ветра с коэффициентом 2.5.
2. Температура в точке становится равна среднему значению температуры солнечного ветра.
3. Скорость в точке становится равна вычисленной скорости КВМ.

Иначе значения всех переменных в этой точке остаются без изменений.

Результаты

В качестве примера для демонстрации работы модели был выбран КВМ за 14 августа 2024 года. Также для произошедшего КВМ были выбраны 8 подряд идущих по времени изображений. Построенные конусы показаны на рисунках 1 и 2. Параметры моделируемого КВМ приведены в таблице 1.

Разница в параметрах одного и того же КВМ объясняется значительными различиями в подходах к их оценке. Разработанная конусная модель использует изображения лишь с одного коронографа LASCO/C3; модель, использованная в каталоге DONKI [8] использует изображения с двух коронографов (LASCO/C3 и STEREO-A/COR2); каталог SEEDS [9] работает без участия человека-оператора.

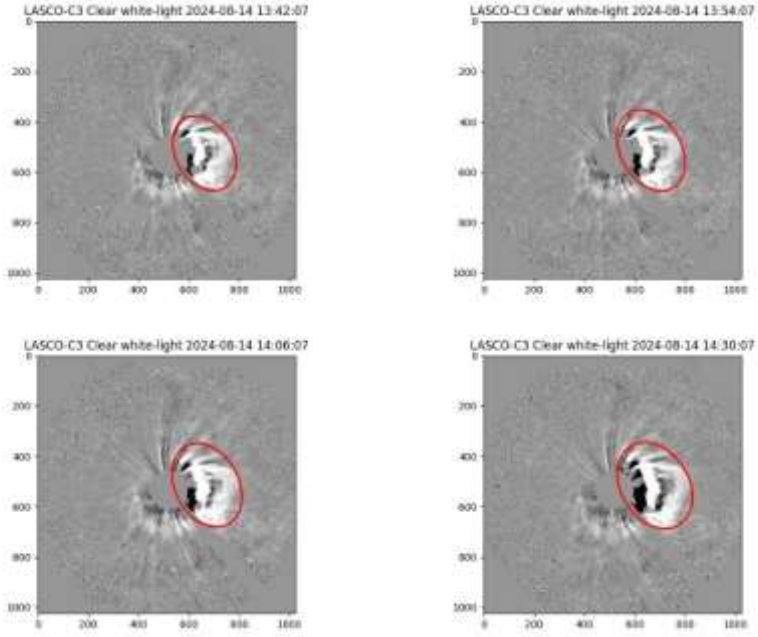


Рис. 1. Построение конусов по первым 4 изображениям коронографа LASCO за 14 августа 2024 года.

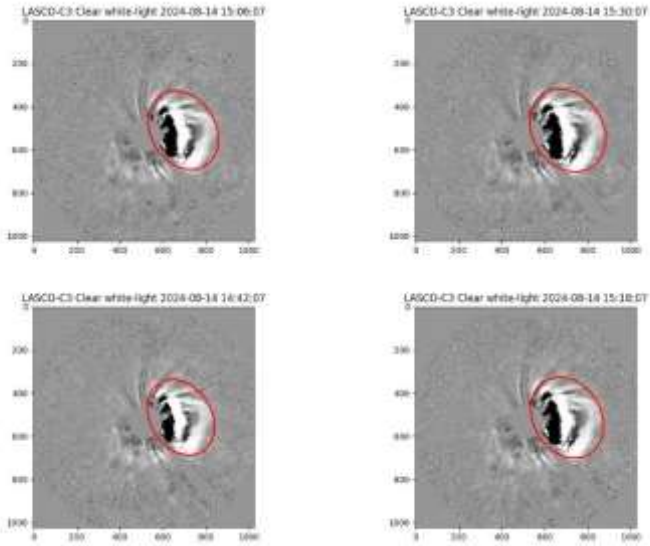


Рис. 2. Построение конусов по последним 4 изображениям коронографа LASCO за 14 августа 2024 года.

Таблица 1

Сравнение результатов работы конусной модели с существующими каталогами

Параметр	Конусная модель	DONKI	SEEDS
Угловой размер (градусы)	66	72	54
Скорость (км/с)	739	426	175
Долгота (градусы)	176	180	-

Широта (градусы)	84	78	-

Эти различия опускаются, поскольку наиболее важным критерием сравнения в космической погоде является время прибытия КВМ к определенной точке в межпланетном пространстве, в основном, это окрестность Земли на расстоянии 1 а.е. от Солнца.

На результатах прогноза солнечного ветра за 14 августа 2024 года наблюдается визуальная схожесть. Более подробное сравнение представлено на рис. 3.

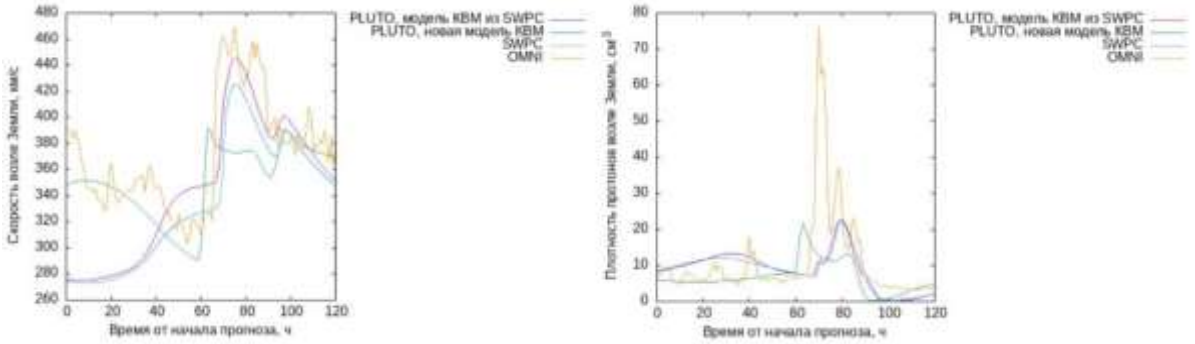


Рис. 3. Сравнение скорости и плотности солнечного ветра и КВМ в окрестности Земли

Заключение

Был разработана конусная модель кинематических и геометрических свойств КВМ. Полученные результаты прогнозирования распространения КВМ в межпланетном пространстве соответствуют зарубежным аналогам и МГД-модели с использованием модели КВМ из SWPC при одинаковых граничных условиях моделирования. Прогноз, полученный при помощи разработанной конусной модели, подобно другим аналогам стремится к пиковым значениям на реальных данных, но с некоторым смещением относительно PLUTO со старой конусной моделью и SWPC. Замеченное смещение возникло из-за неравенства в скорости, полученной при работе конусной модели (в DONKI – 426 км/с, в разработанной модели – 726 км/с). Такая разница в оценке скорости объясняется неточностью, возникающей при оценке скорости на основе изображений, полученных только с одного КА.

В качестве дальнейшего развития работы планируется реализация автоматической системы детектирования КВМ, а также использование данных с нескольких КА.

Список литературы

1. Прогнозы SWPC // WSA-ENLIL Solar Wind Prediction | NOAA / NWS Space Weather Prediction Center [Электронный ресурс]. URL: https://www.swpc.noaa.gov/products/wsa-enlil-solar-wind-prediction (дата обращения: 29.04.2025)

2. Арутюнян С. Н., Кодуков А. В., Субботин М. О., Павлов Д. А. Прототип службы прогноза спокойного солнечного ветра на основе МГД-моделирования и граничных условий модели WSA. Космические исследования. 2023, вып. 61. С. 447-453.

3. Odstrcil D. Modeling 3-D solar wind structure. Advances in Space Research. 2003, iss. 32., pp. 497-506.

4. Веб-сайт PLUTO [Электронный ресурс]. URL: http://plutocode.ph.unito.it/ (дата обращения: 29.04.2025).

5. Арутюнян С. Н. Разработка численной МГД-модели динамики корональных выбросов массы в солнечном ветре // ВКР. 2023, СПбГЭТУ «ЛЭТИ».

6. Domingo V., Fleck B., Poland A. I. The SOHO Mission: an Overview. Solar Physics. 1995, iss. 162, pp. 1-37.

7. Xie H., Ofman L., Lawrence G. Cone model for halo CMEs: Application to space weather forecasting. Journal of Geophysical Research: Space Physics. 2004, iss. 109.

8. Каталог КВМ DONKI [Электронный ресурс]. URL: https://kauai.ccmc.gsfc.nasa.gov/DONKI (дата обращения: 28.04.2025).

9. Каталог KBM SEEDS [Электронный ресурс]. URL: http://spaceweather.gmu.edu/seeds/lasco.php (дата обращения: 20.04.2025).

ТРОПИЧЕСКАЯ КРИПТОГРАФИЯ. РЕАЛИЗАЦИЯ АЛГОРИТМА ЦЭП В ТЕРМИНАХ ТРОПИЧЕСКОЙ АЛГЕБРЫ

ГАВРИЛОВА О.П.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Исследовались возможности использования тропической алгебры для реализации алгоритма цифровой электронной подписи, некоторые варианты атак на алгоритм и его модификация с целью избежания атак рассмотренных типов. На языке программирования Python реализованы алгоритмы цифровой электронной подписи в терминах тропической алгебры, построены графики и многогранники Ньютона получившихся сигнатур.

Ключевые слова: тропическая алгебра, цифровая электронная подпись, криптография.

Введение

В последние годы тропическая математика привлекает всё больше внимания в области криптографии, предлагая принципиально новые подходы к решению задач защиты информации. В отличие от классических алгебраических систем, тропическая алгебра оперирует операциями минимума (или максимума) и сложения, что создаёт нелинейные структуры, устойчивые к ряду криптографических атак.

Актуальность работы обусловлена растущими требованиями к криптостойкости алгоритмов в условиях развития квантовых вычислений. Традиционные схемы цифровых подписей, такие как RSA и ECDSA, уязвимы к алгоритмам Шора и Гровера [1], что стимулирует поиск альтернативных математических моделей. Тропическая алгебра, благодаря сложности решения уравнений в её терминах, представляет собой перспективную основу для создания устойчивых криптографических протоколов. Например, тропические аналоги задач факторизации и дискретного логарифма демонстрируют повышенную вычислительную сложность даже для квантовых компьютеров [2].

Целью данной работы является анализ алгоритма цифровой электронной подписи (ЦЭП), основанного на операциях тропической алгебры и представление результатов в виде графиков и многогранников Ньютона получившихся сигнатур.

Базовые определения

Тропическое полукольцо maxplus-алгебры определяется как

$$\mathbb{T}_{\max} := (\mathbb{R} \cup \{-\infty\}, \oplus, \otimes); x \otimes y := x + y, x \oplus y := \max(x, y), \forall x, y \in \mathbb{R} \cup \{-\infty\}.$$

Тропическое полукольцо minplus-алгебры определяется как

$$\mathbb{T}_{\min} := (\mathbb{R} \cup \{\infty\}, \oplus, \otimes); x \otimes y := x + y, x \oplus y := \min(x, y), \forall x, y \in \mathbb{R} \cup \{\infty\}.$$

Тропический моном представляет собой обычную линейную функцию. Тропический полином – минимум (максимум) из числа таких функций, а, следовательно, является вогнутой кусочно-линейной функций.

Алгоритм ЦЭП

Рассмотрим один из вариантов алгоритма цифровой электронной подписи в терминах тропической алгебры [3].

Закрытый ключ: два многочлена $X, Y \in \mathbb{T}[x]$ степени 2d, с коэффициентами в диапазоне [0, r], где r – один из параметров схемы.

Открытый ключ: $M = X \otimes Y$.

1. Применим хэш-функцию H (SHA-512) к m. $H(m) \to P \in \mathbb{T}[x] : \deg(P) = d$.
2. Выберем два случайных многочлена $U, V \in \mathbb{T}[x] : \deg(U) = \deg(Y), \deg(V) = \deg(X)$ с коэффициентами в диапазоне [0, r].
3. $N = U \otimes V$.
4. Итоговая сигнатура: $(P, P \otimes X \otimes U, P \otimes Y \otimes V, N)$.

На рис. 1 и 2 приведено графическое представление одного из полиномов, входящего в итоговую сигнатуру.

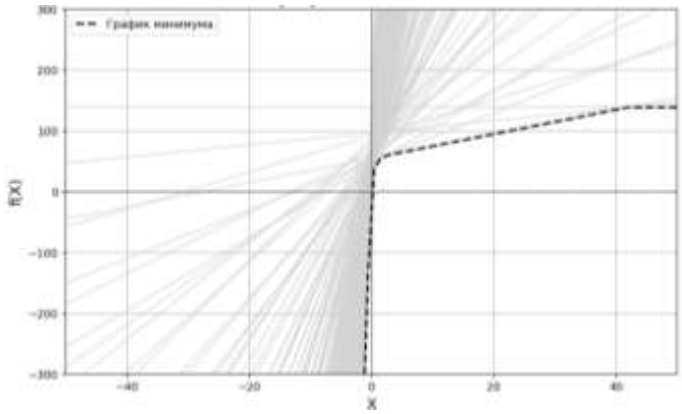


Рис 1. График полинома N, входящего в сигнатуру ЦЭП

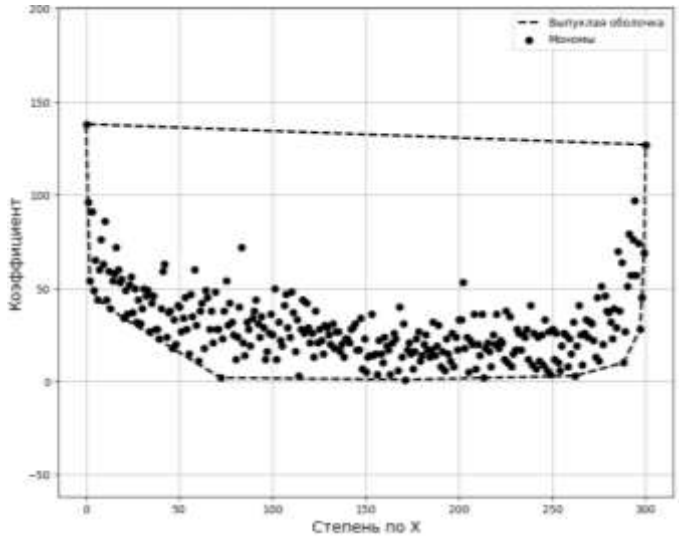


Рис 2. Многогранник Ньютона полинома N, входящего в сигнатуру ЦЭП

Верификация:

1. Вычисление *H(m)* и преобразование *H(m)* в многочлен $P \in \mathbb{T}[x] : \deg(P) = d$. Проверка того, что P – правильный хэш сообщений.
2. Проверка, что степень многочленов $P \otimes X \otimes U$ и $P \otimes Y \otimes V$ равна *2d*, а $\deg(N) = 2d$. Если это не так, то подпись не принимается.

3. Проверка того, что $P \otimes X \otimes U$ и $P \otimes Y \otimes V$ не являются кратными $P \otimes M$ и $P \otimes N$. Если это так, то подпись не принимается.
4. Проверка того, что все коэффициенты в многочленах $P \otimes X \otimes U$ и $P \otimes Y \otimes V$ находятся диапазоне [0, 3r], а все коэффициенты в N находятся в диапазоне [0, 2r].
5. Вычисление $W = (P \otimes X \otimes U) \otimes (P \otimes Y \otimes V)$. Подпись принимается тогда и только тогда, когда $W = P \otimes P \otimes M \otimes N$.

Детерминированная структура и линейные зависимости между компонентами подписи создают уязвимости, упрощающие атаки на факторизацию и подделку данных. Для устранения этих уязвимостей рассмотрим альтернативный алгоритм, ключевым элементом которого является случайный многочлен E [3].

Альтернативный алгоритм ЦЭП

Закрытый ключ: два многочлена $X, Y \in \mathbb{T}[x]$ степени 2d, с коэффициентами в диапазоне [0, r], где r – один из параметров схемы.

Открытый ключ: $M = X \otimes Y$

1. Применим хэш-функцию H (SHA-512) к m. $H(m) \to P \in \mathbb{T}[x] : \deg(P) = 2d$.
2. Выберем два случайных многочлена $U, V \in \mathbb{T}[x] : \deg(U) = \deg(Y), \deg(V) = \deg(X)$ с коэффициентами в диапазоне [0, r].
3. $N = U \otimes V$.
4. Выберем случайный открытый многочлен E степени $3d$ со всеми коэффициентами в диапазоне [0,3r].
5. Итоговая сигнатура: $(P, P \oplus (X \otimes U), P \oplus (Y \otimes V), P \otimes [(X \otimes U) \oplus (Y \otimes V)] \oplus E, N, E)$

Один из полиномов получившейся итоговой сигнатуры изображен на рис. 3 и 4.

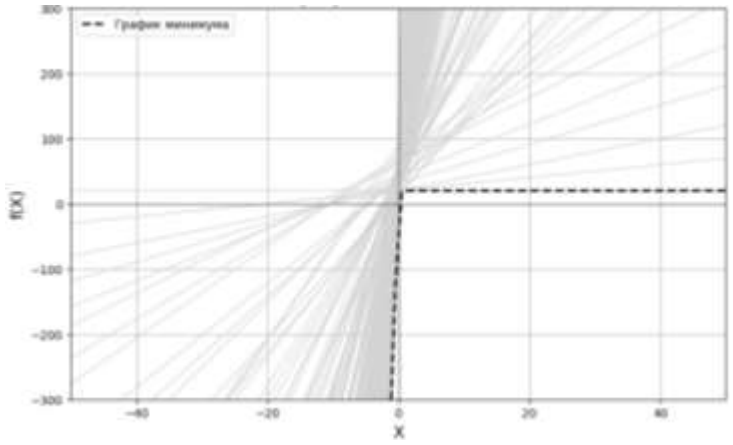


Рис 3. График полинома N, входящего в сигнатуру альтернативной ЦЭП

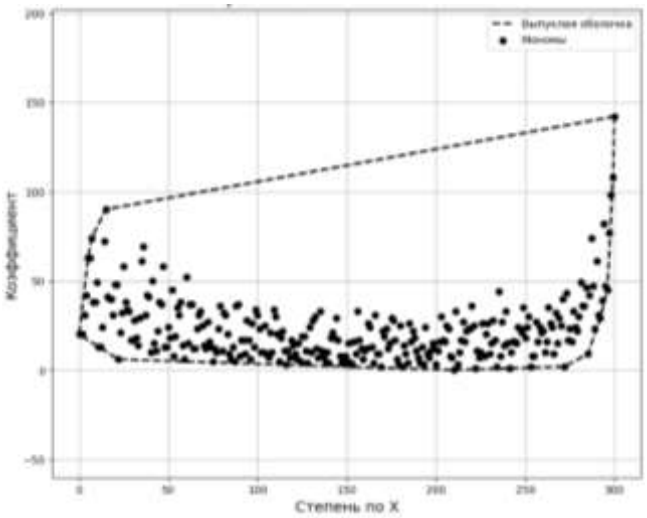


Рис 4. Многогранник Ньютона полинома N, входящего в сигнатуру альтернативной ЦЭП

Верификация:

1. Вычисление *H(m)* и преобразование *H(m)* в многочлен $P \in \mathbb{T}[x] : \deg(P) = 2d.$
2. Проверка, что степень многочленов $P \oplus (X \otimes U)$ и $P \oplus (Y \otimes V)$ равна *2d*, а $\deg(N) = 2d.$ Степени двух оставшихся многочленов равны *3d.* Если это не так, то подпись не принимается.
3. Проверка того, что $P \oplus (X \otimes U)$ и $P \oplus (Y \otimes V)$ не являются кратными $P \oplus M$ и $P \oplus N.$ Если это так, то подпись не принимается.
4. Проверка того, что все коэффициенты в многочленах $P \oplus (X \otimes U)$ и $P \oplus (Y \otimes V)$ находятся в диапазоне [0, 2r], а все коэффициенты в многочлене N находятся в диапазоне [0, 2r]. Все коэффициенты в оставшихся двух многочленов находятся в диапазоне [0, 3r]. Если это не так, то подпись не принимается.
5. Обозначим $R = P \otimes [(X \otimes U) \oplus (Y \otimes V)].$ Проверка, что $P \otimes [P \oplus (X \otimes U)) \oplus (P \oplus (Y \otimes V))] \oplus E = (P \otimes P) \oplus (R \oplus E).$ Если это не так, то подпись не принимается.
6. Вычисление $W = (P \oplus (X \otimes U)) \otimes (P \oplus (Y \otimes V)) = (P \otimes P) \oplus (P \otimes [(X \otimes U) \oplus (Y \otimes V)]) \oplus (X \otimes U \otimes Y \otimes V).$ Подпись принимается тогда и только тогда, когда $W \oplus E = (P \otimes P) \oplus (R \oplus E) \oplus (M \otimes N).$

Результаты

Реализованные алгоритмы обладают временной сложностью $O(k \cdot d^2),$ экспериментальные графики (рис. 5) подтверждают, что альтернативный алгоритм цифровой электронной подписи эффективнее – константа *k* почти в два раза меньше по сравнению с первым представленным алгоритмом ЦЭП. Альтернативный алгоритм увеличивает размер сигнатуры, но является более безопасным, сохраняя при этом квадратичную сложность.

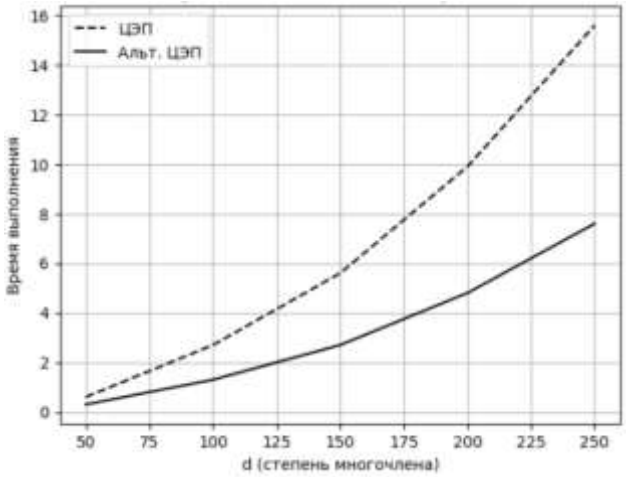


Рис 5. Временная сложность алгоритмов

Исходный код реализованных алгоритмов на языке программирования Python доступен по ссылке [4].

Список литературы

1. Shor, P. W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. — 1997.
2. Grigoriev D., Shpilrain V. Tropical cryptography. — 2018.
3. Chen J., Grigoriev D., Shpilrain V. Tropical cryptography III: digital signature – 2024.
4. Реализованные алгоритмы [Электронный ресурс] https://colab.research.google.com/drive/1EJLW_5AiZzhbtv9SY_Dv9qaArHUidlYH#scrollTo=RPOG7ItqEQov (Дата обращения: 29.04.2025)

ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ ПОДХОДОВ В КЛАССИЧЕСКИХ АЛГОРИТМАХ СТЕРЕОЗРЕНИЯ ПРИ НЕДОСТАТКЕ ИНФОРМАЦИИ

Кодуков А.В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В работе представлена методика поиска и сопоставления опорных точек на изображениях, обладающая повышенной отказоустойчивостью. Рассматривается случай сцены с повторяющимися мало детализированными текстурами, на котором исследуются возможности существующих алгоритмов глубокого обучения для сопоставления изображений. Разработана методика повышения точности и количества опорных точек, продемонстрировано ее преимущество над классическими алгоритмам фотограмметрии.

Ключевые слова: сопоставление опорных точек, техническое зрение, ошибка переноса

Введение

Алгоритмы стереозрения [1] играют ключевую роль в навигации беспилотной техники и трехмерном анализе пересекающихся изображений. Первым и одним из самых проблемных этапов обработки последовательности изображений обычно является попарное сопоставление изображений [2]. Необходимо работать исключительно с пиксельной информацией в условиях изменений ракурса, освещения, масштаба, повторяющихся текстур или мало детализированных объектах. В связи с этим возникают ситуации, в которых изображения даже с достаточно большим пересечением не могут быть сопоставлены, что отменяет всю дальнейшую обработку.

Целью данной работы является создание устойчивой к сложным данным методики поиска соответствий, которая в то же время наиболее равномерно покрывает область пересечения изображений. Для решения поставленной задачи анализируются возможности существующих алгоритмов поиска и сопоставления опорных точек на основе глубокого обучения, рассматривается итеративный подход, комбинирующий особенности различных алгоритмов, и пересматриваются подходы верификации соответствующих точек.

Алгоритмы сопоставления изображений

Обработка пары изображений состоит из поиска опорных точек и их сопоставления между изображениями. Наиболее известные классические алгоритмы: SIFT [3], SURF [4], ORB [5]. Эти алгоритмы извлекают особые точки, пользуясь функцией градиента контрастности, описывают каждую из точек и ее окружение некоторым набором чисел, называемым дескриптором. Дескрипторы совпадающих точек должны быть близки, поэтому в них кодируется поворот, а вычисление точек происходит в нескольких разрешениях. Такие алгоритмы обладают повышенной устойчивостью к аффинным преобразованиям, небольшим изменениям освещенности и локальными изменениями сцены, обладая при этом большой скоростью, что объясняет их использование во всех фотограмметрических программных комплексах, например: COLMAP [6], AgiSoft MetaShape [7], OpenDroneMap [8]. Но классические алгоритмы подходят не во всех случаях, они плохо справляются с изменениями контрастности в снимках разного происхождения, а также с низко детализированными сценами с повторяющимися текстурами.

Для решения этих проблем разрабатываются алгоритмы на основе глубокого обучения, которые лучше используют семантическую информацию изображения. Многие из них не

обрабатывают изображения по отдельности, а принимают на вход сразу пары изображений, что увеличивает время обработки, но повышает шансы найти общие точки в трудных случаях. Однако такие методы намного менее устойчивы к аффинным преобразованиям, потому их часто комбинируют друг с другом и запускают в несколько итераций. Отдельного внимания заслуживают алгоритмы, нацеленные на плотное покрытие области пересечения опорными точками [9], так как для последующей обработки они дают больше всего информации.

После сопоставления точек возникает большое количество ложных срабатываний, что обычно компенсируется геометрической верификацией. Необходимо построить преобразование, переносящее точку с плоскости первого изображения на второе, подходящее для наибольшего количество сопоставленных точек. В качестве такого преобразования используются проективная матрица, подразумевающая, что сцена является плоскостью, которую можно спроецировать на плоскости кадров, определяемые параметрами и положением камер. Для поиска подходящего преобразования в условиях большого количества выбросов используется статистический метод RANSAC [11] или его аналоги.

Метод повышения отказоустойчивости и плотности точек

Было выявлено, что часть алгоритмов более успешно справляется с аффинными преобразованиями, а другая часть более плотно покрывает изображения совпадающими точками, поэтому было решено проводить обработку в 2 этапа: поиск оптимальной проекции и поиск общих точек на одинаково ориентированных снимках. Отбор конкретных алгоритмов приведен в следующем разделе с помощью проекта image-matching-models [12].

После детального осмотра результатов сопоставления пар изображений на конкретном примере аэрофотоснимков крупного здания обнаружено, что фильтрация по матрице проекции может отбросить слишком много правильных соответствий. На изображениях может быть несколько плоскостей на разной высоте (земля и крыша здания), а проективное преобразование подбирает только одну плоскость, отбрасывая точки на других, даже если они не являются реальными выбросами. В связи с этим предлагается строить проективные преобразования по точкам несколько раз, пока количество точек остается выше определенного порога.

Анализ результатов

Тестирование алгоритмов происходило на наборе аэрофотоснимков крыши здания [11]. Вручную были расставлены общие точки на нескольких парах снимков.

В таблице 1 представлены 10 лучших результатов по количеству обработанных пар, а также ошибка переноса ручных точек по матрице проекции, рассчитанной на основе автоматических точек, и доля площади покрытия изображения точками из расчета одна точка – квадрат 10x10. Алгоритм сопоставления точек “в лоб” обозначен в таблице как “NN”, что означает поиск ближайшего соседа.

Таблица 1

10 наиболее отказоустойчивых моделей на отобранных изображениях

Название модели	Ошибка переноса (в пикселях)	Доля обработанных пар	Площадь покрытия
Steerers	47.8	95%	10%
GIM+LightGlue	23.2	95%	18%

DoGHardNet+NN	56.5	83%	11%
Affine Steerers	64.2	79%	16%
SIFT+LightGlue	55.5	79%	10%
SIFT+NN	43.9	75%	10%
DoGHardNet+LightGlue	149.4	70%	12%
RoMa	40.9	62%	32%
minima-RoMa	29.9	54%	37%

Заметно, что наиболее отказоустойчивые алгоритмы используют только самые надежные точки, из-за чего покрывают только небольшую часть изображения. Можно отметить, что в список вошёл классический алгоритм SIFT, что свидетельствует о его высокой надёжности. Также заметно преимущество модели LightGlue [13] среди алгоритмов, не имеющих собственного метода сопоставления точек. Для этого этапа выбран алгоритм GIM [14]+LightGlue, как обладающей наименьшей ошибкой переноса.

Затем изображения были спроецированы под каждую из своих пар по опорным точкам алгоритма GIM+LightGlue, чтобы увеличить визуальную схожесть. После повторного запуска всех алгоритмов, только 3 обработали все пары снимков (таблица 2), из которых RoMa обладает наибольшей площадью покрытия.

Таблица 2

Анализ сопоставленных точек на полученных алгоритмом GIM+LightGlue изображениях

Название модели	Площадь покрытия	Количество опорных точек	Ошибка переноса (в пикселях)
Matchformer	34.9%	1883	7.84
RoMa	42%	1196	7.851
LoFTR	35.3%	1643	7.853

Заметно, что итоговая ошибка переноса значительно уменьшилась относительно первого запуска. Пример работы представлен на рисунке 1: сопоставление неориентированных изображений алгоритмом GIM+LightGlue, проектирование первого изображения на плоскость второго, вторая итерация сопоставления изображений алгоритмом RoMa.

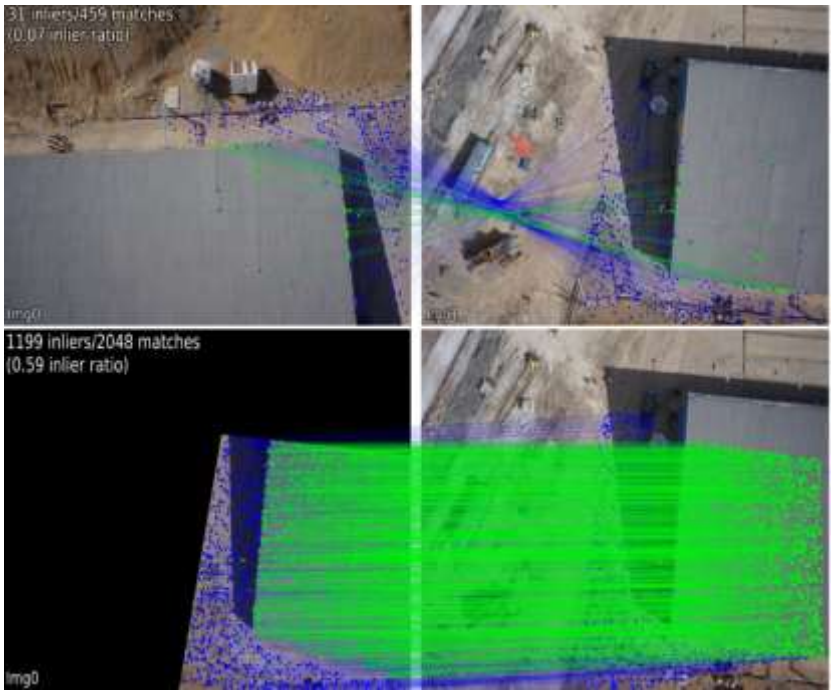


Рис. 1. Сопоставление неориентированных изображений алгоритмом GIM+LightGlue (сверху), вторая итерация сопоставления изображений алгоритмом RoMa на одинаково ориентированных изображениях (снизу)

На каждом этапе матрица проекции строится 2 раза, после чего выбирается та из них, где ошибка переноса ручных точек меньше. Таким образом, вычисления гарантированно относятся к плоскости крыши, где и расставлены ручные точки.

С помощью разработанной методики для данного набора изображений относительно обработки обычным SIFT удалось увеличить количество срабатываний в 1.3 раза, уменьшить ошибку переноса в 5.6 раз, увеличить площадь покрытия в 4.2 раза.

Список литературы

1. Hartley R., Zisserman A. Multiple view geometry in computer vision. – Cambridge university press, 2003.
2. Ma J. et al. Image matching from handcrafted to deep features: A survey //International Journal of Computer Vision. – 2021. – Т. 129. – №. 1. – С. 23-79.
3. Lowe G. Sift-the scale invariant feature transform //Int. J. – 2004. – Т. 2. – №. 91-110. – С. 2.
4. Bay H., Tuytelaars T., Van Gool L. Surf: Speeded up robust features //Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9. – Springer Berlin Heidelberg, 2006. – С. 404-417.
5. Rublee E. et al. ORB: An efficient alternative to SIFT or SURF //2011 International conference on computer vision. – Ieee, 2011. – С. 2564-2571.
6. Schonberger J. L., Frahm J. M. Structure-from-motion revisited //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 4104-4113.
7. Agisoft Metashape [Электронный ресурс]: программа для фотограмметрической обработки изображений / Agisoft LLC. – Санкт-Петербург: Agisoft, 2024. – Режим доступа: https://www.agisoft.com/ (дата обращения: 28.04.2025).
8. OpenDroneMap (ODM) [Электронный ресурс] : программное обеспечение для фотограмметрической обработки аэрофотоснимков с БПЛА / OpenDroneMap Community. – 2020. – Режим доступа: https://github.com/OpenDroneMap/ODM (дата обращения: 29.04.2025).
9. Edstedt J. et al. RoMa: Robust dense feature matching //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2024. – С. 19790-19800.
10. Fischler M. A., Bolles R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography //Communications of the ACM. – 1981. – Т. 24. – №. 6. – С. 381-395.

11. Яндекс.Диск [Электронный ресурс]: архив изображений для обработки и результатов / – Режим доступа: https://disk.yandex.ru/d/2-bIXXB4RM7KEw (дата обращения: 27.04.2025).
12. Berton G. et al. Earthmatch: Iterative coregistration for fine-grained localization of astronaut photography //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2024. – С. 4264-4274.
13. Lindenberger P., Sarlin P. E., Pollefeys M. Lightglue: Local feature matching at light speed //Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2023. – С. 17627-17638.
14. Shen X. et al. Gim: Learning generalizable image matcher from internet videos //arXiv preprint arXiv:2402.11095. – 2024.

ПОИСК ЦЕНТРАЛЬНЫХ МЕР НА ГРАФАХ ДВУСТРОЧЕЧНЫХ И ТРЕХСТРОЧЕЧНЫХ ДИАГРАММ ЮНГА

КРУГЛОВА В. Д.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Реализован быстрый алгоритм поиска центральной меры на произвольном конечном градуированном графе, содержащем сток. С помощью алгоритма найдены оценки переходных вероятностей для графа двустрочечных диаграмм Юнга. Установлено, что свойство однородности выполняется с высокой точностью, причем оценки вероятностей перехода близки к вычисленным теоретически. Результаты исследования на графе двустрочечных диаграмм Юнга показали, что реализованный алгоритм может быть применен для поиска центральных мер и на других градуированных графах, в том числе на графе трехстрочечных диаграмм Юнга.

Ключевые слова: центральная мера, граф двустрочечных диаграмм Юнга, граф трехстрочечных диаграмм Юнга, асимптотическая комбинаторика

Введение

На градуированном графе можно задать процесс случайного блуждания или процесс Маркова. В этом случае каждому ребру будет сопоставлен вес, равный вероятности перехода по этому ребру. Таким образом на графе задается вероятностная мера. Данная мера называется центральной, если для любой фиксированной пары вершин вероятности перехода по двум разным путям между ними равны. Поиск центральных мер на градуированных графах является актуальной проблемой асимптотической комбинаторики [1]. Существуют различные способы поиска центральных мер на градуированных графах [2, 3], в данной работе используется т. н. алгоритм на основе вычисления коразмерностей.

Граф двустрочечных диаграмм Юнга – бесконечный градуированный граф, вершины которого являются диаграммами Юнга, количество клеток в столбцах которых не превышает 2, а каждое ребро соединяет диаграмму λ_i *i*-го этажа с диаграммой λ_{i+1}, полученную добавлением одной клетки к λ_i. На рисунке 1 изображено начало данного графа, каждое ребро отмечено соответствующей переходной вероятностью μ_n, $1 - \mu_n$ или 1.

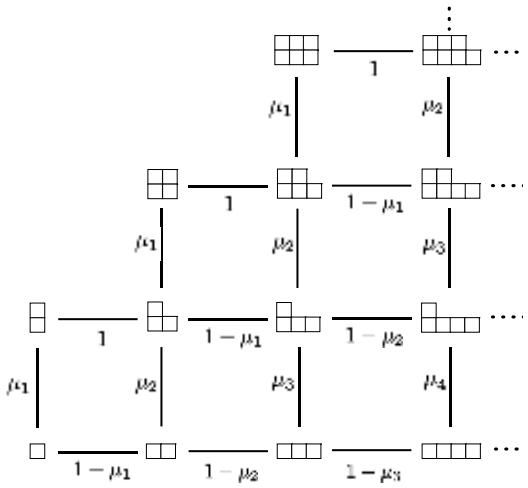


Рис. 1 – начало графа двустрочечных диаграмм Юнга

На графе двустрочечных диаграмм Юнга может выполняться свойство однородности. Мера является однородной, если при добавлении слева к паре диаграмм, соединенных ребром, произвольного количества столбцов максимальной для рассматриваемого графа высоты, вероятность перехода по ребру, соединяющему полученную пару диаграмм, будет равна исходной переходной вероятности.

На рассматриваемом графе существует единственная однородная центральная мера, для которой вероятности перехода по ребрам стремятся к 1/2 с ростом *n,* и для этой меры $\mu_1 = 1/4$ [4]. Соответствующие ей переходные вероятности могут быть вычислены по формуле:

$$\mu_n = \frac{1}{2} - \frac{1}{2(n+1)}. \quad (1)$$

На графе трехстрочечных диаграмм Юнга также может выполняться свойство однородности. Для этого графа не существует формулы, аналогичной (1), которая бы описывала однородную центральную меру с похожими асимптотическими свойствами. Можно предположить, что с помощью выбранного алгоритма можно получить оценки соответствующих переходных вероятностей.

Быстрый алгоритм поиска центральной меры на градуированном графе

Алгоритм основан на поиске копереходных вероятностей «перевернутого» графа, в котором корнем является сток исходного графа, а ребра ориентированы в противоположном направлении. Копереходные вероятности для любого градуированного графа со стоком можно вычислить единственным образом как отношение количества путей, ведущих из корня в каждую из вершин, соединенных ребром. Копереходные вероятности для «перевернутого» графа эквивалентны переходным вероятностям исходного графа, на котором задана центральная мера. Тогда копереходные вероятности будут вычисляться как отношение коразмерностей – количеств путей в каждую из вершин, соединенных ребром, из стока исходного графа (корня «перевернутого» графа). Алгоритм на основе вычисления коразмерностей состоит из следующих шагов:

1. Каждой вершине назначается некоторый вес (коразмерность). Вес корневой вершины полагается равным 1.
2. Оставшиеся вершины обрабатываются последовательно вплоть до стока:
   2.1 в качестве веса для вершины сохраняется количество входящих путей – складываются веса всех вершин, соединенных с ней входящим ребром;
   2.2 для каждого входящего ребра записывается вероятность перехода по нему как отношение весов вершин, соединенных этим ребром.

Таким образом, необходимо всего один раз обойти граф, чтобы получить искомые значения оценок переходных вероятностей.

Анализ полученных данных для графа двустрочечных диаграмм Юнга

Рассмотрим подграф, состоящий из первой 1000 этажей графа двустрочечных диаграмм Юнга, при этом каждая вершина последнего этажа соединена единственным исходящим ребром с абстрактной вершиной (стоком). На этом графе с помощью алгоритма на основе вычисления коразмерностей можно получить оценки вероятностей перехода центрального марковского процесса и определить, к какой именно центральной мере удалось прийти (определить значение μ_1). Также необходимо определить, насколько точно выполняется свойство однородности.

В результате работы алгоритма получены оценки переходных вероятностей для ребер, соединяющих диаграмму, состоящую из единственной строки, с диаграммой, образованной добавлением клетки в первый ее столбец. В таблице 1 приведены оценки переходных вероятностей для некоторых из этих ребер и вероятности перехода из формулы (1) для них.

Таблица 1

Сравнение теоретических и полученных в ходе исследования переходных вероятностей на некоторых этажах графа двустрочечных диаграмм Юнга

μ_n	Ребро	Теоретически полученная вероятность	Практически полученная вероятность	Ошибка оценки
μ_1	{1} → {2}	0.25	0.25025	0.00025
μ_2	{1, 1} → {2, 1}	0.(333)	0.333778	0.000445
μ_3	{1, 1, 1} → {2, 1, 1}	0.375	0.375626	0.000626
μ_4	{1, 1, 1, 1} → {2, 1, 1, 1}	0.4	0.400801	0.000801
μ_{38}	{1, 1, …, 1} → {2, 1, …, 1}	0.486842	0.492628	0.005786
μ_{74}	{1, 1, …, 1} → {2, 1, …, 1}	0.493243	0.499364	0.006121

Как видно из таблицы, разница между теоретическими и эмпирическими значениями переходных вероятностей искомой центральной меры ($\mu_1 = 1/4$), не превышает 10^{-2} для первых 74 этажей. При этом вероятности перехода по рассматриваемым ребрам стремятся к 1/2 с увеличением n. Для того, чтобы выяснить, насколько точно выполняется свойство однородности, обратимся к таблице 2, где представлены эмпирические оценки вероятностей перехода по ребрам, для которых вероятность перехода должна быть одинаковая.

Таблица 2

Проверка выполнимости условия однородности

μ_4 Этаж ребра	Теоретическая вероятность	Разница с предыдущим значением μ_4
4	0.400801	----------
6	0.400803	0.000002
8	0.400804	0.000001
10	0.400806	0.000002
12	0.400807	0.000001
14	0.400809	0.000002
70	0.400858	----------
72	0.400859	0.000001
74	0.400861	0.000002

Можно заметить, что разница первого встреченного значения μ_4 и того, которое было вычислено на 74 этаже, меньше ошибки оценки для ребер из таблицы 1. Таким образом, при использовании алгоритма на основе вычисления коразмерностей для графа двустрочечных диаграмм Юнга свойство однородности выполняется с достаточно высокой точностью.

Заключение

Исследовались переходные вероятности центрального марковского процесса на графе, состоящем из тысячи первых этажей графа двустрочечных диаграмм Юнга, где каждая из диаграмм последнего этажа соединена единственным исходящим ребром со стоком. Для поиска центральной меры на исследуемом конечном графе использовался алгоритм на основе вычисления коразмерностей. В результате проведенного вычислительного эксперимента было обнаружено, что алгоритм может быть применен для оценки вероятностей перехода по ребрам бесконечного графа двустрочечных диаграмм Юнга, на котором задана однородная центральная мера, при которой вероятность перехода по ребрам стремится к 1/2 с увеличением количества клеток в диаграммах. При этом ошибка оценки не превышает 10^{-2} для ребер первых 74 этажей графа. Также установлено, что на исследуемом конечном графе свойство однородности выполняется с высокой точностью.

Можно предположить, что с помощью реализованного алгоритма на графе трехстрочечных диаграмм Юнга можно найти однородную центральную меру, также обладающую асимптотическими свойствами.

Литература

1. Thoma E. Die unzerlegbaren, positiv-definiten Klass- enfunktionen der abzählbar unendlichen, symmetrischen Gruppe. Mathematische Zeitschrift, 1964, vol. 85, pp. 40–61.
2. Vershik A. M., Kerov S. V. Asymptotic of the largest and the typical dimensions of irreducible representations of a symmetric group. Functional Analysis and its Applications, 1985, vol. 19, iss. 1, pp. 21–31. doi:10.1007/BF01086021
3. Круглова В.Д. Поиск центральных мер на конечных градуированных графах, ННБ XII, Санкт-Петербург, 2024, Том 1, стр. 267-271.
4. Вершик А.М. Спектр и абсолют графа двустрочечных диаграмм Юнга, 2022

РАЗРАБОТКА МЕТОДОВ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ СТАТИСТИЧЕСКОЙ ОЦЕНКИ СЕЛЕКТИВНОСТИ ЗАПРОСОВ В РЕЛЯЦИОННЫХ СУБД С ИСПОЛЬЗОВАНИЕМ КОРРЕЛЯЦИОННОГО АНАЛИЗА

НАЙДЁНОВА Е. В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В работе представлен метод повышения точности оценки селективности SQL-запросов в реляционных СУБД на основе корреляционного анализа данных. Разработан подход протягивания матрицы корреляций через дерево операторов запроса. Для оценки операций агрегации (GROUP BY) разработан подход на основе копул. Этот метод позволяет точно оценивать количество уникальных групп, формируемых при группировке, с учетом степени корреляции между столбцами

Ключевые слова: оценка селективности, оптимизация запросов, матрица корреляций, копулы, GROUP BY

Введение

Эффективность выполнения SQL-запросов зависит от точности оценки селективности операторов — доли кортежей, удовлетворяющих заданному предикату — на основе которой оптимизатор формирует план запроса [1]. Традиционное допущение о независимости атрибутов приводит к значительным ошибкам при наличии корреляций между столбцами, что влечет выбор неэффективных алгоритмов обработки данных, превращая потенциально быстрые запросы в ресурсоемкие операции [2].

Обзор существующих подходов

Анализ существующих методов повышения точности оценки селективности выявил их основные ограничения:

- Отсутствие динамического учета изменения корреляций в процессе выполнения запроса;
- Зависимость от накопления исторических данных запросов;
- Ограниченная применимость только к числовым атрибутам;
- Использование эмпирических зависимостей без строгого статистического обоснования.

Предлагаемый метод

Разработанный метод основан на использовании матрицы корреляций, которая протягивается по дереву операторов запроса с последовательной модификацией при прохождении каждого оператора. Метод включает два ключевых компонента:

1. Формирование и модификация матрицы корреляций при прохождении операторов;
2. Уточнение селективности операторов с использованием информации о корреляциях.

Формирование исходной матрицы корреляций

Предварительно для каждой таблицы в системном каталоге СУБД сохраняются значения корреляции Cramer's V для всех пар столбцов. Cramer's V является обобщением корреляции Пирсона и в случае применения к числовым переменным его значение совпадает с абсолютным значением коэффициента корреляции Пирсона. Коэффициент Крамера вычисляется по формуле:

$$V = \sqrt{\frac{\chi^2/n}{\min(d_1-1, d_2-1)}} \quad \chi^2 = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\frac{(n_{ij}-E_{ij})^2}{E_{ij}} \ ,$$

где:

- n — общее количество наблюдений (записей) в выборке
- d_1 — количество уникальных значений первой переменной
- d_2 — количество уникальных значений второй переменной
- n_{ij} — наблюдаемая частота совместного появления
- $n_{i.}$ — общее количество записей с i-ым значением первой переменной
- $n_{.j}$ — общее количество записей с j-ым значением второй переменной

При построении плана запроса оптимизатор использует предварительно рассчитанные матрицы корреляций для таблиц и динамически модифицирует их с учетом применяемых операторов запроса.

Оценка кардинальности агрегации с использованием корреляции.

Разработан метод для более точной оценки числа уникальных групп при выполнении GROUP BY по двум столбцам A и B, который учитывает корреляцию между ними.

Ожидаемое число уникальных групп оценивается по формуле:

$$d_1 \cdot d_2 - \sum_{a=0}^{d_1-1}\sum_{b=0}^{d_2-1}(1-p_{a,b})^N \ .$$

Для оценки $p_{a,b}$ используется копула [3]. Копула имеет следующую функцию распределения:

$$C_\alpha(u,v) = \min(\min(u,v), uv + \alpha u(1-u)v(1-v)), \alpha \geq 0 \ .$$

Для получения дискретных вероятностей $p_{j,i}$ для каждой ячейки сетки размера $d_1 \times d_2$, квадрат $[0,1] \times [0,1]$ разбивается на прямоугольники:

$$\left[\frac{j}{d_1}, \frac{j+1}{d_1}\right) \times \left[\frac{i}{d_2}, \frac{i+1}{d_2}\right), j = 0, \dots, d_1 - 1, i = 0, \dots, d_2 - 1 \ .$$

Вероятность ячейки (j,i) вычисляется как:

$$p_{j,i} = C_\alpha\left(\frac{j+1}{d_1}, \frac{i+1}{d_2}\right) - C_\alpha\left(\frac{j}{d_1}, \frac{i+1}{d_2}\right) - C_\alpha\left(\frac{j+1}{d_1}, \frac{i}{d_2}\right) + C_\alpha\left(\frac{j}{d_1}, \frac{i}{d_2}\right) .$$

Эта формула является дискретным аналогом плотности распределения и получается из вычисления смешанной второй производной функции $C_\alpha(u,v)$.

При $|\alpha| > 1$ в определённых областях квадрата вероятность становится равной нулю. Для учёта этого при дискретизации выделяются граничные области копулы.

Чтобы упростить последующее суммирование, удобно задать область $\mathcal{R}$, в пределах которой $\mathrm{p_{j,i}}$ принимается равным нулю. Для ячеек в области $\overline{\mathcal{R}}$, то есть лежащих в области с ненулевой вероятностью, $\mathrm{p_{j,i}}$ рассчитывается по общей разностной формуле копулы. Таким образом, итоговое двойное суммирование было разбито на две части.

Тогда итоговую сумму можно записать так:

$$\mathrm{d_1 \cdot d_2} - \underbrace{\sum\sum_{\mathrm{(j,i)}\in\mathcal{R}}(1-0)^{\mathrm{N}}}_{\text{ячейки,}\mathrm{p_{j,i}=0}} - \underbrace{\sum\sum_{\mathrm{(j,i)}\in\overline{\mathcal{R}}}(1-\mathrm{p_{j,i}})^{\mathrm{N}}}_{\text{ячейки,}\mathrm{p_{j,i}>0}} .$$

В области $\mathcal{R}$, так как $p_{j,i} = 0$, часть $(1-0)^N$ равна 1, и сумма сводится к числу ячеек в этой области. Кроме того, при $\alpha > 1$ копула симметрична относительно диагоналей квадрата, поэтому нужно рассчитывать только четверть области с ненулевой вероятностью.

Экспериментальные результаты

Данные в каждом столбце таблицы 1 были сгенерированы равномерно на соответствующих доменах, затем вычислялась корреляция данных, на основе которой проводилась оценка числа групп, и полученное значение сравнивалось с реальным количеством групп в результирующем наборе данных. Результаты, представленные в таблице, демонстрируют высокую точность разработанного метода при оценке числа уникальных групп в различных сценариях.

Таблица 1

Оценка числа уникальных групп

Параметры	1	2	3	4	5
N (число кортежей)	5000	10000	10000	1000	10000
d_1	90	50	36	226	100
d_2	90	50	29	491	100
$d_1 \cdot d_2$	8100	2500	1044	110966	10000
Корреляция Cramer's V	1.0000	1.0000	0.0047	0.0383	1.0000
Параметр α	118.1004	118.1004	0.0140	0.1150	118.1004
Фактическое число групп	90	50	1043	996	100
Расчёт по формуле	90.0000	50.0000	1043.9280	995.5055	100.0000
Погрешность	0.0000%	0.0000%	0.0069%	0.0495%	0.0000%

Заключение

Разработан метод повышения точности оценки селективности запросов на основе корреляционного анализа, который решает следующие задачи :

1. Обеспечивает динамический учет изменения корреляций при прохождении операторов в дереве запроса;
2. Использует метод на основе копул для точной оценки кардинальности агрегации;
3. Применим к различным типам данных, не ограничиваясь числовыми атрибутами.

Список литературы

1. Jarke M., Koch J. Query optimization in database systems //ACM Computing surveys (CsUR). – 1984. – Т. 16. – №. 2. – С. 111-152.
2. Ilyas I. F. et al. CORDS: Automatic discovery of correlations and soft functional dependencies //Proceedings of the 2004 ACM SIGMOD international conference on Management of data. – 2004. – С. 647-658.
3. Augustyn D. R. Copula-based module for selectivity estimation of multidimensional range queries //Man-Machine Interactions 5: 5th International Conference on Man-Machine Interactions, ICMMI 2017 Held at Kraków, Poland, October 3-6, 2017. – Springer International Publishing, 2018. – С. 569-580.

ВЫЧИСЛЕНИЕ ОРБИТЫ КА MESSENGER ПО РАДИОТЕХНИЧЕСКИМ ИЗМЕРЕНИЯМ ДОПЛЕРОВСКОГО СДВИГА ЧАСТОТЫ В X-ДИАПАЗОНЕ

СУББОТИН М.О.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Точные межпланетные дальности необходимы для построения современных планетных эфемерид. В частности, дальности Земля-Меркурий позволяют не только определять орбиту Меркурия, но и изучать физические свойства Солнца и проверять общую теорию относительности. Существуют публично доступные точки дальности, полученные в JPL из необработанных радиотехнических данных MESSENGER, однако в данной работе будет представлен независимый анализ. В работе были обработаны исходные радиотехнические данные космического аппарата (КА) MESSENGER. Проведено интегрирование орбиты КА MESSENGER с учетом неоднородного гравитационного потенциала Меркурия и влияния Солнца, Луны и планет. Затем орбита КА MESSENGER была уточнена путем ее подгонки к данным радионаблюдения (доплеровский сдвиг частоты).

Ключевые слова: эфемериды, Меркурий, MESSENGER, доплеровский сдвиг частоты, интегрирование орбиты

Введение

В современных исследованиях солнечной системы важную роль играют высокоточные астрометрические наблюдения, в частности радиотехнические наблюдения орбитальных аппаратов, позволяющие построить точные эфемериды Солнца и планет. Эфемериды служат уникальным инструментом для экспериментальной проверки общей теории относительности [1] и для изучения гравитационных свойств Солнца. Настоящая работа, в сочетании с отечественными исследованиями последних лет, способствует уточнению ранее полученных оценок [2] и может привести к качественно новым научным результатам.

Цель и задачи

Цель работы заключается в вычислении и уточнении орбиты космического аппарата (КА) MESSENGER. Для выполнения поставленной цели следует выполнить следующие задачи:

- обработка радиотехнических наблюдений: доплеровского смещения частоты с учетом калибровочных поправок и релятивистской модели распространения сигнала;
- численное интегрирование уравнений движения КА MESSENGER с учетом неоднородного гравитационного потенциала Меркурия и возмущений, оказываемых Солнцем, Луной и планетами;
- нахождение параметров орбиты КА MESSENGER по наблюдениям доплеровского смещения частоты.

Обработка радиотехнических наблюдений

КА MESSENGER [3] был запущен в космос 3 августа 2004 года с авиационной станции Кейп-Канаверал, Флорида, США. Это был первый КА, который вышел на орбиту Меркурия. Полет изначально предполагался на 8 лет, но миссия несколько раз продлевалась и закончилась в апреле 2015 года. Одной из целью КА MESSENGER было получение радионаблюдений.

Наблюдаемые величины представлены в файлах орбитальных данных (ODF, Orbit Data File) [4]. ODF является стандартным бинарным файлом формата JPL (Jet Propulsion Laboratory) и содержит множество 36-байтовых записей, которые делятся на 4 основные группы. Для чтения ODF была использована библиотека pds4-tools [5], написанная на языке Python 3.

Рассматриваемая величина, а именно доплеровский сдвиг частоты, может быть:

односторонней – DSN (Deep Space Network) станция «А» отправила сигнал в момент времени t1 и КА получил сигнал в момент времени t2;

двухсторонней – DSN станция «А» отправила сигнал в момент времени t1, КА получил сигнал в момент времени t2 и отправил его обратно DSN станции «А», станция «А» приняла сигнал в момент времени t3;

трехсторонней – DSN станция «А» отправила сигнал в момент времени t1, КА получил сигнал в момент времени t2 и отправил его DSN станции «Б», станция «Б» приняла сигнал в момент времени t3.

Доплеровские наблюдаемые величины получаются из изменения количества доплеровских циклов N, которые накапливаются за время сжатия Tc = tj - ti на принимающей станции. Формула величины наблюдаемого доплеровского смещения частоты:

$$Obs = \left(\frac{B}{|B|}\right) \times \left[\left(\frac{N_j - N_i}{t_j - t_i}\right) - |F_b \times K + B|\right],$$

где B – смещение на приемнике; ti и tj – время начала и конца интервала; Ni, Nj – количество доплеровских циклов в момент времени ti, tj; Fb – смещение частоты; K – коэффициент усиления сигнала (turnaround ratio), для X-диапазона коэффициент K равен 11/3. Значение Fb для двухстороннего и трёхстороннего доплеровского сдвига частоты:

$$F_b = (X_1/X_2) \times (X_3 \times f_{qR} + X_4) \; - \; (T_1/T_2) \times (T_3 \times f_{qT} + T_4),$$

где fqR – частота генератора приемника в момент времени t3; fqT – частота генератора передатчика в момент времени t1; T1, …, T4 – константы, зависящие от диапазона; X1, …, X4 – константы, зависящие от диапазона частоты возбуждения, значения аналогичные T1, …, T4.

Вычисление модельного доплеровского смещения частоты

Формула для модельного двухстороннего и трёхстороннего доплеровского смещения частоты следующая:

$$F_{2,3} = M_2 f_T(t_3) - \frac{M_2}{T_c} \int_{t_{1s}}^{t_{1e}} f_T(t_1) dt_1, \qquad (1)$$

где M2 – коэффициент усиления сигнала (turnaround ratio), fT – частота передатчика, t1s и t1e – время начала и конца интервала сжатия. Интеграл в (1) вычисляется по процедуре из [6].

Динамическая модель движения КА MESSENGER

Для интегрирования динамической модели использовался интегратор ABMD [7]. Динамическая модель содержит как гравитационные силы, так и негравитационные. Гравитационные силы включают:

- гравитационный потенциал Меркурия;
- твердотельные приливы;
- возмущения, вызванные Солнцем, Луной и планетами;

- релятивистские поправки.

Негравитационные силы включают давление солнечного излучения и тепловое излучение. Для учета негравитационных сил использовалась макро-модель «Box-Wing» космического корабля [8], которая представляет КА в виде куба (основное тело) и двух солнечных панелей по бокам.

Решение обратной задачи

Для подгонки вычисленных модельных значений доплеровского смещения частоты к данным радионаблюдения использовался метод наименьших квадратов. В методе наименьших квадратов свободными параметрами выступали начальные значения положения и скорости КА MESSENGER. На рисунке 1 представлены графики невязки (разность модельной и наблюдаемой величин) доплеровского сдвига частоты для орбиты КА MESSENGER из JPL Horizons [9], орбиты без уточнения и орбиты после уточнения начальных значений.

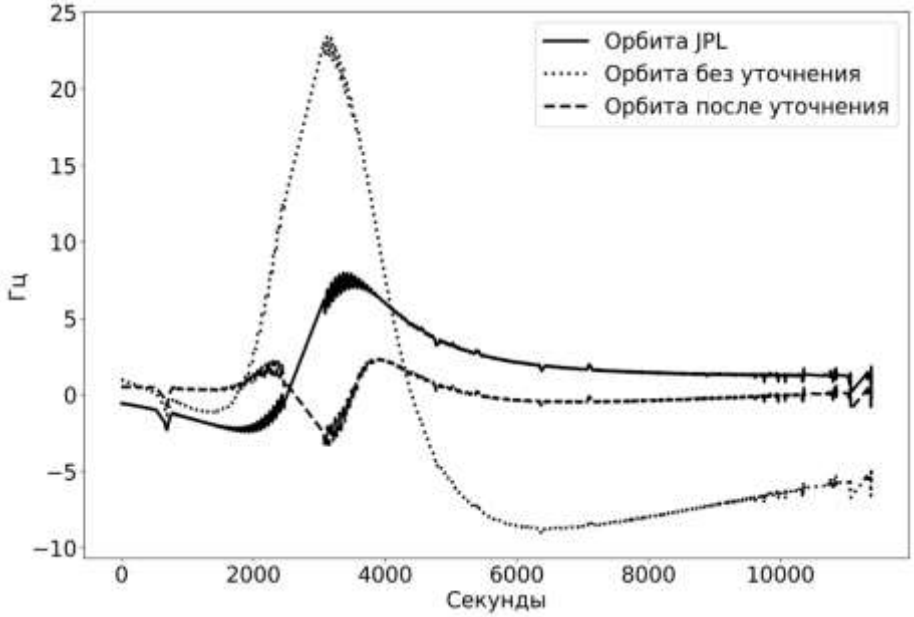


Рис. 1. График невязки доплеровского сдвига частоты (DSN №26)

Можно заметить, что амплитуда невязки доплеровского сдвига частоты орбиты после уточнения меньше по сравнению с амплитудой орбиты из JPL Horizons.

Результаты

В работе решалась задача моделирования орбиты КА MESSENGER и уточнения ее начальных значений. Обработаны исходные данные радионаблюдения (доплеровская задержка частоты). Написана программа по интегрированию орбиты КА MESSENGER и подгонки вычисленных величин к данным радионаблюдения. В дальнейшем будут проведены вычисления точек дальности Меркурия. Также развитие работы заключается в изучении динамических свойств Солнца и фундаментальных свойств пространства-времени.

Список литературы

1. Pavlov D., Dolgakov I. (2024). Studying the Properties of Spacetime with an Improved Dynamical Model of the Inner Solar System. Universe, 10(11), 413. doi:10.3390/universe10110413
2. Park R. S., Folkner W. M., Konopliv A. S., Williams J. G., Smith D. E., Zuber M. T. (2017). Precession of Mercury's Perihelion from Ranging to the MESSENGER Spacecraft. The Astronomical Journal, 153(3), 121. doi:10.3847/1538-3881/aa5be2
3. Srinivasan D. K., Perry M. E., Fielhauer K. B. et al. The Radio Frequency Subsystem and Radio Science on the MESSENGER Mission. Space Sci Rev 131, 557–571 (2007). doi:10.1007/s11214-007-9270-7
4. Verma A. K., Perry M. E., Lemoine F. G., Smith D. E., Ensor S., Reid M.R. MESSENGER Radio Science Orbit Data File (ODF) Product Collection [Электронный ресурс]. URL: https://pds-ppi.igpp.ucla.edu/data/mess-rs-raw/data-odf (Дата обращения: 28.04.2025)
5. Nagdimunov L. Package to read and display NASA PDS4 data [Электронный ресурс]. URL: https://pypi.org/project/pds4-tools/ (Дата обращения: 29.04.2025)
6. Moyer T. D. 2003, Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation (John Wiley & Sons), 2
7. Aksim D., Pavlov D. On the Extension of Adams–Bashforth–Moulton Methods for Numerical Integration of Delay Differential Equations and Application to the Moon’s Orbit. Math.Comput.Sci. 14, 103–109 (2020). doi: 10.1007/s11786-019-00447-y
8. Marshall J. A., 1992. Modeling radiation forces acting on TOPEX/Poseidon for precision orbit determination.
9. Giorgini J. D. JPL Solar System Dynamics Group, "NASA/JPL Horizons On-Line Ephemeris System." [Электронный ресурс] URL: https://ssd.jpl.nasa.gov/horizons. (Дата обращения: 03.05.2025)

СЕКЦИЯ ИНФОРМАТИКА И УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ И ВТ

РАЗРАБОТКА АЛГОРИТМА ЭФФЕКТИВНОЙ ЦИФРОВИЗАЦИИ ПРОФИЛАКТИЧЕСКИХ ОСМОТРОВ В ВЕДОМСТВЕННЫХ ПОЛИКЛИНИКАХ МВД И МЧС

Вздорова М.А.

Санкт-Петербургский государственный экономический университет

Аннотация. В статье рассмотрена проблематика применения тиражируемых медицинских информационных систем (МИС), имеющихся на отечественном рынке, для цифровизации ежегодного профилактического осмотра сотрудников МВД и МЧС в ведомственных поликлиниках на территории Российской Федерации, а также предложения по решению данной проблемы.

Ключевые слова: информатизация, цифровизация медицины, ведомственные медицинские учреждения, профилактический осмотр, экономическая эффективность, проблемы информатизации государственных учреждений.

Основная часть. На текущий момент уровень цифровизации ведомственных поликлиник МВД и МЧС в ряде случаев отстает от аналогичных общегражданских государственных и частных медицинских учреждений, подведомственных Министерству здравоохранения Российской Федерации [1]. Это отставание обусловлено рядом факторов, в числе которых различные регламенты проведения ежегодных профилактических осмотров, различия в структуре и формах медицинской статистики. Соответственно, в ведомственных поликлиниках невозможно прямое внедрение типовых конфигураций, тиражируемых МИС, а также использование готовых модулей обмена данными между МИС и ЕГИСЗ (Единой государственной информационной системой в сфере здравоохранения) и другими смежными информационными системами.

Для сравнительного анализа процесса ежегодного профилактического медицинского осмотра было выполнено детальное сравнение соответствующих приказов МВД и МЧС РФ [2, 3]. Результаты сравнений обязательных объемов обследований в виде таблицы представлены на рисунке 1.

Обязательные врачи специалисты	
Гражданские	**Ведомственные**
Гинеколог	Гинеколог
Невролог	Невролог
Оториноларинголог	Оториноларинголог
Офтальмолог	Офтальмолог
Стоматолог	Стоматолог
Терапевт	Терапевт
Уролог	
Хирург	Хирург
	Психиатр
Исследования в рамках осмотра	
Гражданские	**Ведомственные**
ЛАК	ЛАК
	Глюкоза холестерин
Измерение внутриглазного давления	Измерение внутриглазного давления
КАК	КАК
Кал на скрытую кровь	Кал на скрытую кровь
Кровь на ЗППП	
Маммография / PSA	Маммография / PSA
Общий анализ мочи	Общий анализ мочи
	Спирография
	УЗИ брюшной полости
	УЗИ органов малого таза
	Узи почек
Флюорография	Флюорография
	ХТИ
Цитологический цервикальный мазок	Цитологический цервикальный мазок
ЭКГ	ЭКГ

Рис. 1 Сравнение объемов обследований взрослого гражданского населения и сотрудников ведомственных структур

При проведении сравнительного анализа выявлено, что существуют значительные различия в проведении осмотра гражданского населения и сотрудников ведомственных служб. Разница существенна уже на первом этапе осмотра, а именно в способе получения информации по датам и возможности пройти профилактический осмотр. Различны также состав врачей-специалистов, которые участвуют в осмотре, количество анализов, которые должны сдавать сотрудники ведомственных структур и виды исследований, которые с ними проводятся. Это обусловлено важностью своевременной диагностики отклонений от норм в показателях здоровья на ранних этапах и выявления факторов риска хронических заболеваний, чтобы врачи могли своевременно взять под контроль развивающиеся заболевания и начать их профилактику, а пациенты продолжить службу.

Также в отличии от гражданского населения, для сотрудников структурных подразделений МВД и МЧС в ходе осмотра необходимо определять уровень физической подготовки. Это делается врачом-терапевтом и другими врачами-специалистами путем сбора нескольких показателей, например, силы сжатия кисти.

Дополнительными ограничениями являются повышенные требования к защите информации от несанкционированного доступа, предъявляемые в ведомственных медицинских учреждениях МВД и МЧС к разработке, внедрению и эксплуатации программных средств. Например, большинству программных средств, предлагаемых рынком для профилактических осмотров, для работы необходимо подключение к интернету, что запрещено. Кроме того, в большинстве ведомственных поликлиник используется устаревшая компьютерная техника, не соответствующая техническим требованиям к установке современных МИС. Её замена требует значительных финансовых вложений, а выделяемый на цифровизацию бюджет в текущих реалиях ограничен.

Также невозможно и создание единого сервера, который бы позволил масштабировать одну адаптированную МИС на все поликлиники. По требованиям безопасности [4] на одну поликлинику приходится одна сеть интранет, которая содержит базу данных, информацию с которой в электронном виде за пределы поликлиники передавать нельзя.

Разработка медицинской информационной системы "с нуля" требует внимание Государства, так как бюджет проектов подобного уровня сложности стартует от восьми миллионов рублей без четко обозначенного верхнего предела.

Несмотря на трудности цифровизации любых медицинских работ, многие из них поддаются строгим алгоритмам и могут быть автоматизированы. Изучая данные из Приложение № 3 к приказу МВД России от 24.04.2019 № 275 [2] был разработан алгоритм формирования направления лечения в виде схемы.

На рисунке 2 представлена часть алгоритма создания направления лечения в рамках ЕМПО.

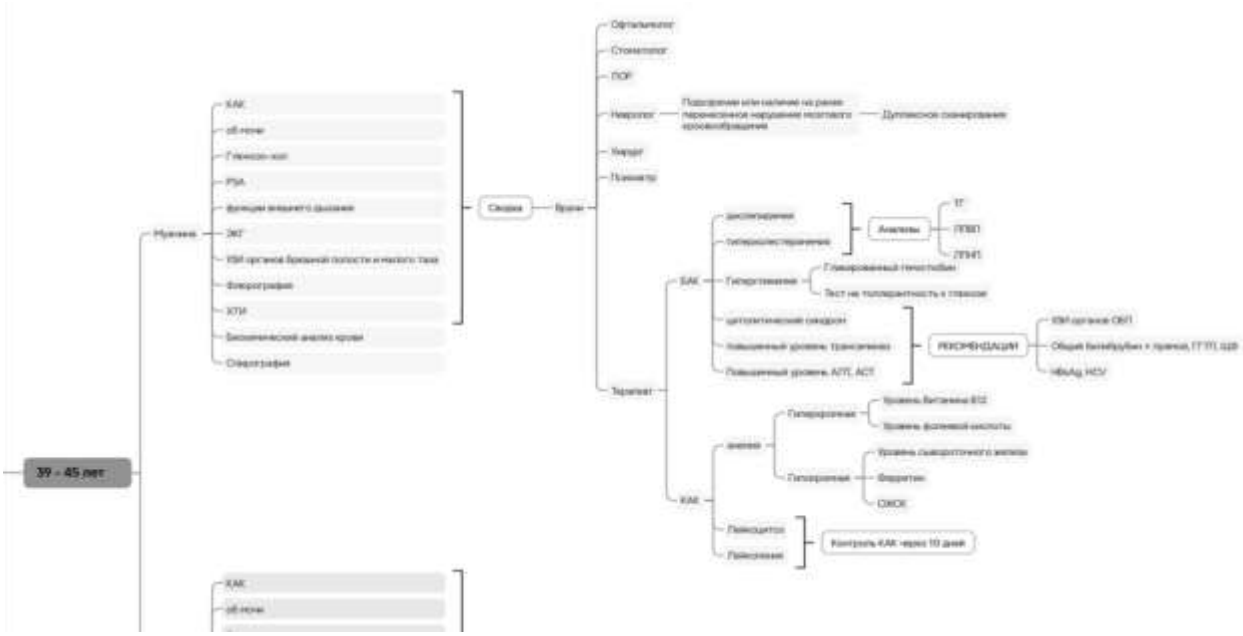


Рис. 2 Часть алгоритма создания направления лечения в рамках ЕМПО

В процессе изучения процесса профилактического осмотра и создания алгоритма создается понимание, что процессы, как например, составление лечения практически невозможно автоматизировать, так как на решение, принимаемое врачом, влияет огромное количество факторов. Однако, процессы, как например, разработка рекомендаций и выписка дополнительных направлений при нахождений в анализах отклонения от нормы поддаются чётким инструкциям. В таблице 1 в качестве примера на основе интервью с врачами ведомственного учреждения собраны отклонения от нормы в клиническом анализе крови и рекомендации, которые должна выдавать система.

Таблица 1

Отклонения от нормы клинического анализа крови и рекомендации, которые должна выдавать система

Наименование анализа	Отклонение от нормы	Необходимые дообследования
КАК	гипохромная микроцитарная анемия	уровень сывороточного железа, ферритин, ОЖСК
	гиперхромная макроцитарная анемия	уровень витамина B12, уровень фолиевой кислоты
	лейкопения	контроль КАК через 10 дней , консультация гематолога

	лейкоцитоз	контроль КАК через 10 дней, консультация гематолога
	тромбоцитопения	контроль КАК через 10 дней, консультация гематолога
	тромбоцитоз	контроль КАК через 10 дней, консультация гематолога
	ускоренное СОЭ	контроль КАК через 10 дней

Помимо подсказки в заключениях система для профилактического осмотра должна автоматизировать расчёт следующих медицинских параметров: сердечно-сосудистый риск, индекс массы тела, индекс курильщика, группа состояния физической подготовки и т.д. Так как все они рассчитываются по давно известным формулам.

Ограничения безопасности и бюджета не позволяют создать единую МИС для поликлиник МВД и МЧС, но точечная цифровизация возможна — через автоматизацию анализа медицинских данных и ввод электронного документооборота. Это снизит нагрузку на врачей и подготовит почву для будущего внедрения цифровых инструментов.

Список литературы

1. Мельник О. И. Проблемы ИТ в здравоохранении: кто виноват и что делать? [Электронный ресурс]. 2020. URL: https://www.itworld.ru/cionews/business/157564.html (дата обращения: 20.09.2024).
2. Порядок прохождения профилактических медицинских осмотров сотрудниками органов внутренних дел Российской Федерации, включающих в себя химико-токсикологические исследования наличия в организме человека наркотических средств, психотропных веществ и их метаболитов : Приложение № 3 к приказу МВД России от 24.04.2019 № 275.
3. Приказ Министерства здравоохранения РФ от 27 апреля 2021 г. N 404н "Об утверждении Порядка проведения профилактического медицинского осмотра и диспансеризации определенных групп взрослого населения" (с изменениями и дополнениями).
4. Инструкция по организации защиты персональных данных, содержащихся в информационных системах органов внутренних дел Российской Федерации: Приложение к приказу МВД РФ от 6 июля 2012 г. № 678.

ТЕХНИЧЕСКИЕ АСПЕКТЫ СТЕНДА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ПОЖАРНОЙ СИГНАЛИЗАЦИИ, КОНТРОЛЯ ЗАГАЗОВАННОСТИ И ПОЖАРОТУШЕНИЯ

ГЕНЕРАЛОВ Г. Ю.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Аннотация. В статье рассматривается техническая реализация стенда автоматизированной системы пожарной сигнализации, контроля загазованности и пожаротушения (АСПС, КЗиПТ). Основное внимание уделено аппаратным компонентам системы, включая программируемые логические контроллеры (ПЛК), сетевую инфраструктуру, устройства преобразования интерфейсов и модули управления. Описаны принципы работы системы, ее архитектура и ключевые технические решения.

Ключевые слова: промышленная автоматизация, автоматизированная система пожарной сигнализации (АСПС), контроль загазованности и пожаротушение (КЗиПТ), программируемый логический контроллер (ПЛК)

Автоматизированная система пожарной сигнализации, контроля загазованности и пожаротушения (АСПС, КЗиПТ) является критически важным элементом обеспечения безопасности на промышленных объектах, в жилых зданиях и общественных местах. В современных условиях особую актуальность приобретает развитие отечественных решений в области промышленной автоматизации. Современные системы промышленной безопасности требуют комплексного подхода к проектированию и тестированию. Разработанный стенд автоматизированной системы пожарной сигнализации, контроля загазованности и пожаротушения представляет собой полнофункциональную систему автоматизированного управления (САУ), выполненную на базе отечественных компонентов.

На стенде реализованы несколько конфигураций, а именно 2 независимых контроллера отечественных фирм: «ТРЭИ» и «Текон», две различные схемы пожарных извещателей и обработка сигналов от них:

1. физические имитаторы с использованием модуля охранно-пожарной сигнализации (МОПС-03) и модуля управления пожарной сигнализацией (МУПС-03)
2. цифровой имитатор датчиков, работающих по протоколу R3-Link, с использованием контроллера адресных устройств (КАУ-2).

Для отображения визуальной информации используется панель местного управления фирмы «Weintek», для каждого контроллера используется своя персональная SCADA система, установленная на инженерном ноутбуке.

Характеристики ПЛК «ТРЭИ» [1]:

1. Четырех ядерный микропроцессор
2. Отдельный контроллер ST-BUS, обеспечивающий транспортный протокол внутренней сети при обмене данными с модулями ввода/вывода.
3. ПЗУ объемом 4 ГБ
4. ОЗУ объемом 256 МБ
5. Встроенные энергонезависимые часы реального времени предназначены для выдачи информации о времени и календарной дате
6. 4 дополнительно устанавливаемых съемных юнита (UNIT 1-4), для которых есть клеммы подключения на передней части модуля.

Основные характеристики ПЛК «Текон» [2]:

1. 32-битный процессор с тактовой частотой 533 МГц
2. ПЗУ объемом 512МБ
3. ОЗУ объемом 256МБ
4. Встроенные энергонезависимые часы реального времени предназначены для выдачи информации о времени и календарной дате

Для объединения всех устройств в единую сеть используются отечественные промышленные коммутаторы фирмы АО «ТРЭИ», главными особенностями которых являются:

- поддержка скорости передачи данных по портам Ethernet до 1000 Мбит/с и порту SFP-10000 Мбит/с на большие расстояния;
- степень защиты IP 20, прочный металлический корпус;
- поддержка диагностического доступа по протоколам SNMPv2 и SNMPv3 (Simple Network Management Protocol);

- поддержка сбора диагностических данных через протокол SYSLOG (UDP);
- простые средства конфигурации на основе браузера. Коммутатор может быть легко сконфигурирован по сети с использованием WEB-интерфейса.

Также для передачи данных по протоколу Modbus используется сервер последовательных интерфейсов S341, который позволяет обеспечить связь с программируемого логического контроллера с несколькими устройствами, производящими обмен данными по интерфейсам RS-232/485.

Модуль охранно-пожарной сигнализации рассчитан на подключение шлейфовых линий с токовыми неадресуемыми пожарными извещателими, обеспечивающих или рост тока в шлейфе при срабатывании, или уменьшении тока при срабатывании.

Конфигурация шлейфов хранится в энергонезависимой памяти. Они могут быть считаны или изменены по интерфейсу RS-485. Каждый шлейф имеет следующие настройки:

1. стратегия работы;
2. значения четырех порогов тока;
3. значения четырех таймеров.

Модуль управления пожарной сигнализацией предназначен для управления исполнительными механизмами автоматизированных систем пожарной сигнализации, контроля загазованности и пожаротушения.

Модуль получает команды по интерфейсу RS-485, преобразует его в управляющее воздействие на исполнительные механизмы. Он имеет четыре контролируемые цепи управления исполнительными механизмами.

Каждая контролируемая цепь исполнительного механизма может находиться в одном из четырех состояний: «Норма выключен», «Норма включен», «Обрыв», «КЗ»

Исполнение поданной по интерфейсу команды включения исполнительного механизма возможно только при нахождении соответствующей контролируемой цепи в состоянии «Норма».

Контроллер адресных устройств «Рубеж» предназначен для сбора данных о состоянии адресных пожарных извещателей на адресных линиях связи, передачи этих данных по запросу через интерфейс RS-485, а также приема команд по интерфейсу RS-485 и передачи их адресным пожарным извещателям.

Контроллер, относится к комбинированным устройствам, способным обмениваться с пожарными извещателями и другими техническими средствами пожарной автоматики как пороговыми, так и аналоговыми данными в виде оцифрованных данных.

Разработанный стенд позволяет обучить персонал работать с разными конфигурациями автоматизированной системы пожарной сигнализации, контроля загазованности и пожаротушения. Данный стенд так же позволяет производить отладку алгоритмов работы с полевыми устройствами (извещателями и оповещателями), производить тестирование нового оборудования, работающего по стандартизированным протоколам.

Список литературы

1. УСТРОЙСТВО ПРОГРАММНОГО УПРАВЛЕНИЯ TREI-5B. TREI-5B-05 модули серии Smart TP. Руководство по эксплуатации TREI.421457.001-07 РЭ.
2. МНОГОФУНКЦИОНАЛЬНЫЙ КОНТРОЛЛЕР МФК1500. Руководство по эксплуатации. Часть 1. Общие сведения о контроллере ДАРЦ.420002.003РЭ1.

РЕАЛИЗАЦИЯ БИТ-ПОТОКОВОГО ПЕРСЕПТРОНА НА ПЛИС

ЗДЕХ С.А.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»им. В.И. Ульянова (Ленина)

Аннотация. В работе реализован бит-потоковый персептрон на ПЛИС с использованием пяти различных схем умножения потоков. Рассмотрены архитектуры преобразователей, проведён сравнительный анализ по ресурсоёмкости, частоте и качеству выходного сигнала. Реализации протестированы на задаче бинарной классификации, показана их работоспособность и пригодность для использования в данном варианте персептрона.

Ключевые слова: персептрон, бит-потоковый преобразователь, ПЛИС

1. Введение

Персептрон – это математическая модель нейрона головного мозга, простейшая единица нейронной сети [1]. Он представляет собой бинарный классификатор, предсказания которого основаны на линейной комбинации входных значений:

$$y = \begin{cases} 1 & \text{if } \sum_{i-1}^{n} w_i x_i + b > 0 \\ 0 & \text{if } \sum_{i-1}^{n} w_i x_i + b \leq 0 \end{cases}$$

где y – выходное значение, вектор x – набор входных значений и вектор w – вектор весов, соответствующих каждому входному значению. Сумма произведений входных значений и весов перед подачей на выход поступает в функцию активации, которая обычно имеет вид единичной ступенчатой функции. На рисунке 1 представлена базовая модель персептрона.

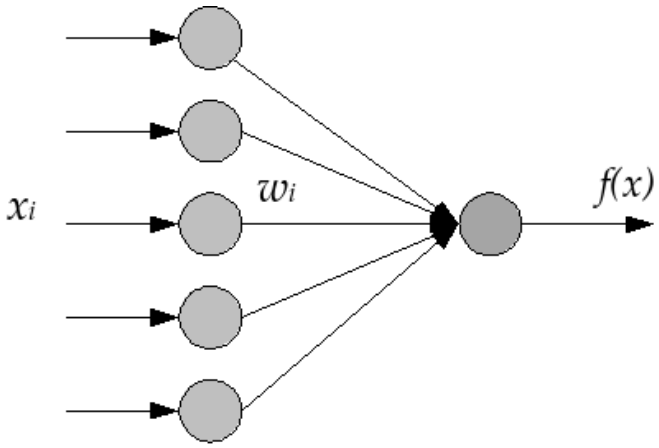


Рис. 1. Персептрон

Бит-потоковый преобразователь [2] или умножитель потоков – это цифровое устройство, цель которого заключается в изменении исходной частоты и выводе полученной последовательности импульсов в течение того же периода времени.

Частота входного потока импульсов f_{in} уменьшается в соответствии с коэффициентом умножения K, который может находиться в пределах от 0 до $2^n - 1$, где n - разрядность коэффициента. Частота выходного потока:

$$f_{\text{out}} = \text{f}_{\text{in}} \times \frac{K}{2^n - 1},$$

то есть, на выходе умножителя мы получаем последовательность из K импульсов за один период, а сам коэффициент задается на входе схемы в виде двоичного числа.

Бит-потоковый персептрон представляет собой реализацию нейронной сети, в которой данные и вычисления представлены в виде последовательностей импульсов. В отличие от традиционных цифровых систем, оперирующих двоичными числами, он использует количество импульсов в потоке за определенный период для представления числовых значений. Это позволяет значительно упростить аппаратную реализацию арифметических операций, таких как умножение и сложение, выполняя их с помощью простых логических элементов: И, ИЛИ и реверсивных двоичных счетчиков.

В данной архитектуре входные сигналы, представленные в виде потоков импульсов, умножаются с помощью бит-потоковых преобразователей на весовые коэффициенты, закодированные в двоичной форме. Результирующие потоки затем суммируются с учетом знаков весов с использованием сумматоров, преобразующих плотность импульсов в цифровое значение.

2. Бит-потоковые преобразователи

Для реализации и тестирования работы бит-потокового персептрона были выбраны пять архитектур умножителей потоков, которые различаются принципом работы, числом составных элементов, наличием выбросов, максимальной частотой работы и видом выходного потока. Использованы преобразователи: 1 – на основе Т-триггера и одного счетчика, 2 – с использованием набора D-триггеров и одного счетчика [3], 3 – с двумя счетчиками [4], 4 – с одним счетчиком [5], 5 – на основе сумматора [6]. В таблице 1 приведено сравнение результатов синтеза этих преобразователей при разрядности n коэффициента K равной 8. Значения частоты были получены с использованием модели slow 1200 mV 85 °C, предоставляемой инструментом TimeQuest Timing Analyzer среды Quartus II.

Таблица 1

Сравнение результатов синтеза преобразователей

№ Преобразователя		1	2	3	4	5
Всего логических элементов		17	17	25	15	17
Всего комбинационных функций		17	11	25	15	17
Логические элементы по числу входов	4	8	2	5	5	0
	3	1	1	6	0	14
	≤2	8	8	14	10	3
Логические элементы по режиму работы	нормальный	11	5	13	8	3
	арифметический	6	6	12	7	14
Регистры		9	16	16	8	8
Максимальная частота, МГц		528	555	550	547	506
Наличие выбросов при симуляции		есть	нет	нет	есть	есть

Преобразователи 2 и 3, которые при симуляции не давали выбросов, являются одновременно наиболее затратными по используемым ресурсам ПЛИС и самыми быстрыми. Пятый преобразователь, в котором не используются логические элементы с числом входов больше двух, оказался самым медленным. Четвертый умножитель потоков суммарно использует меньше всего ресурсов и показывает достаточно высокую скорость работы. Все пять преобразователей будут использованы при построении бит-потокового персептрона.

3. Бит-потоковый персептрон

Чтобы проверить, способен ли перцептрон на основе умножителей потоков классифицировать поступающую к нему на вход информацию, мы построим простую сеть, задачей которой будет отличать изображение знака “+” от изображения знака “-”. Для обучения этой сети и дальнейшего тестирования персептрона искусственно создан специальный датасет, состоящий из одноканальных изображений размером 3*3 пикселя. Он подается на вход простой нейронной сети, состоящей из полносвязного слоя Linear layer [7], который преобразует 9 входных значений в одно выходное. Значение на выходе, близкое к нулю, будет обозначать класс «+», близкое к единице – класс «-». Обучить эту сеть было необходимо для получения вектора весов w для реализации персептрона.

На рисунке 2 показано графическое представление схемы бит-потокового персептрона, полученное после компиляции проекта в Quartus II. В ней для удобства вектор весов и вектор входных значений были заданы константными, поэтому входных сигнала только два – clock и reset.

Для того, чтобы представить значения яркости пикселей (вектор x) в форме потоков импульсов, одинаковый тактовый сигнал подается на линию input_line[].inst_in, которая состоит из преобразователей код-частота. Этот сигнал умножается на двоичное значение яркости, и мы получаем входной поток импульсов персептрона.

Далее потоки «пикселей» в зависимости от их положения на изображении и знака соответствующего веса из вектора w подаются на вход линии элементов-умножителей потоков input_line[].inst_neg или input_line[].inst_pos, где происходит умножение на коэффициент, обозначающий вес пикселя.

Последняя часть персептрона – «функция активации», которая состоит из трёх элементов: bc_pos и bc_neg, которые считают соответственно число импульсов пришедшее в текущий такт от входов с положительными и отрицательными весами, и final_sum, в котором происходит прибавление этих чисел к текущему значению сумматора. Поскольку начальное значение на сумматоре равно нулю, мы можем использовать его старший бит в качестве указателя класса изображения: если это «+», то оно будет отрицательным (старший бит равен единице), а если «-», то положительным (старший бит равен нулю).

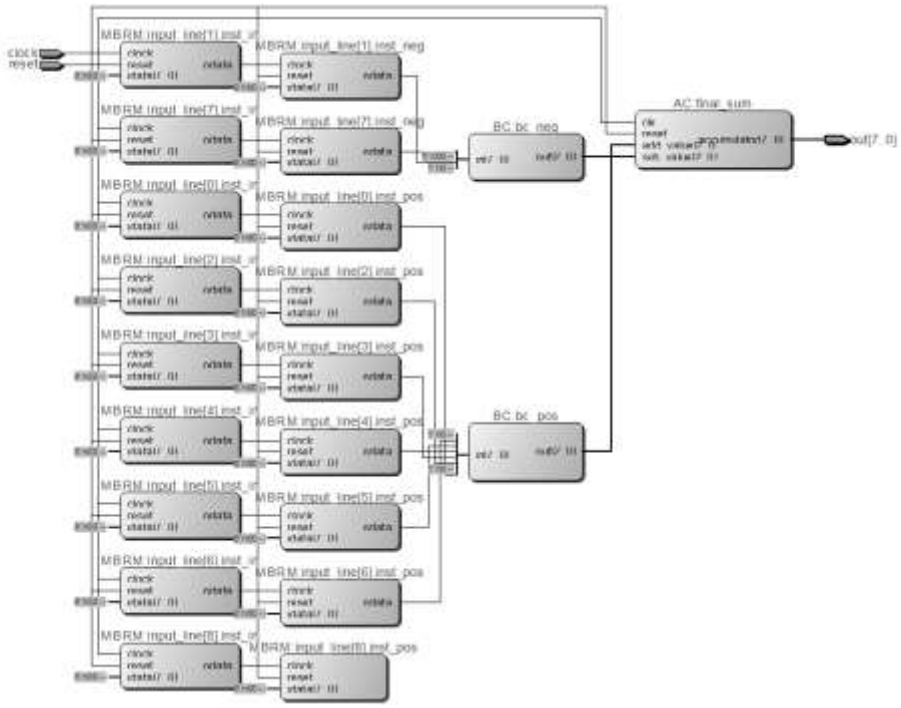


Рис. 2. Результат компиляции бит-потокового персептрона в Quartus II

4. Сравнение персептронов, использующих разные бит-потоковые преобразователи

Для того, чтобы сравнить пять персептронов, будем использовать два изображения и набор весов, которые задаются следующим образом:

{8'd0, 8'd3, 8'd2, 8'd120, 8'd115, 8'd127, 8'd0, 8'd10, 8'd25} – изображение «-»;

{8'd0, 8'd120, 8'd2, 8'd120, 8'd115, 8'd127, 8'd34, 8'd115, 8'd25} – изображение «+»;

{8'd55, 8'd91, 8'd2, 8'd15, 8'd25, 8'd91, 8'd0, 8'd127, 8'd49} – вектор весов.

Начальное значение сумматора, по старшему биту которого мы определяем принадлежность изображения одному из классов, равно нулю, т.е. соответствует классу «плюс», поэтому время установления значения класса можно определить только для изображения с минусом, что представлено в таблице 2. Номер персептрона соответствует номеру преобразователя из первой таблицы.

Таблица 2

Время установления значения для разных персептронов

№ Преобразователя	1	2	3	4	5
Время, нс	440	50	—	60	1760

Схема на основе третьего преобразователя оказалась неработоспособной, тогда как оставшиеся четыре варианта показали разное время работы. Пятый преобразователь оказался самым долгим, в то время как второй и четвертый показали очень малое время работы. Несмотря на то, что импульсы 2 и 4 преобразователя равномерно распределялись по всему периоду, в отличие от первого преобразователя, у которого они концентрировались в начале периода, время установки значения класса у него выше.

5. Заключение

В ходе работы был реализован и исследован бит-потоковый персептрон на ПЛИС с использованием пяти различных схем умножения потоков, которые различаются скоростью, ресурсоёмкостью, видом и стабильностью выходного сигнала. В ходе экспериментов было обнаружено, что преобразователь номер 3 не подходит для использования в составе персептрона. Лучшие результаты показали преобразователи 2 и 4, которые значительно различаются по числу используемых ресурсов ПЛИС.

Список литературы

1.What is a Perceptron? What is the role of bias in a perceptron (or neuron)? [Электронный ресурс]. – Режим доступа: https://aiml.com/what-is-a-perceptron/ (дата обращения: 8.05.2025).
2. Binary Rate Multipliers [Электронный ресурс]. – Режим доступа: https://circuitcellar.com/resources/quickbits/binary-rate-multipliers/ (дата обращения: 8.05.2025).
3. Bureneva O.I. Element Base of Pulse Neural Network Units // Proceedings of the II International Conference on Neural Networks and Neurotechnologies (NeuroNT), Saint Petersburg, Russian Federation. – 2021. – DOI: 10.1109/NeuroNT53022.2021.9472789.
4. Bureneva O., Safyannikov N., Mironov S. Symmetry in the Bit-Stream Converter Design // Symmetry. – 2023. – Vol. 15, No. 11. – DOI: 10.3390/sym15112006.
5. Vasantha S. High speed binary rate multiplier // Proceedings of the Indian Academy of Sciences. Section A. – 1967. – Vol. 65, No. 6. – P. 340–348. – DOI: 10.1007/BF03054095.
6. Oberman R.M.M. A Flexible Rate Multiplier Circuit with Uniform Pulse Distribution Outputs // IEEE Transactions on Computers. – 1972. – Vol. C-21, No. 8. – P. 896–899. – DOI: 10.1109/TC.1972.5009048.
7. PyTorch documentation [Электронный ресурс]. – Режим доступа: https://pytorch.org/docs/stable/index.html (дата обращения: 8.05.2025).

КОНТРОЛЛЕР МОДУЛЯ ФОТОБИОМОДУЛЯЦИИ

Ковальчук И.А.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В рамках проекта, связанного с контролем и коррекцией психофизиологического состояния человека, необходимо обеспечить стимуляцию определенных участков мозга излучением с определенной длиной волны на основе результатов мониторинга показателей электроэнцефалограммы (ЭЭГ). Для этого разработан контроллер беспроводного управления светоизлучающим диодом (СИД) с помощью широтно-импульсной модуляции (ШИМ). В статье показан процесс разработки: обоснован выбор микроконтроллера; представлены серверная часть – приложение на компьютер, разработанное на языке C++ с помощью библиотеки Qt, и клиентская часть – программное обеспечение контроллера. Показаны результаты тестирования: успешная беспроводная передача данных и процесс управления ФБМ. Разработано ПО для анализа ЭЭГ, выполняющее обработку сигналов ЭЭГ, их вывод на дисплей, а также инструмент управления отображений сигналов для детального анализа.

Ключевые слова: фотобиомодуляция, мониторинг ЭЭГ, беспроводная передача данных, C++, Qt, ESP32-C3

В рамках проекта, связанного с контролем состояния человека, решаются задачи беспроводного управления светоизлучающими диодами (СИД) с помощью широтно-импульсной модуляции (ШИМ), необходимые для стимуляции определенных участков мозга, на основе мониторинга показателей ЭЭГ. Для решения задачи управления диодами необходимо разработать контроллер, приложение для компьютера и обеспечить беспроводную стабильную передачу данных между ними.

Для построения такой системы была выбрана клиент-серверная архитектура [1], ориентированная на беспроводную передачу данных по интерфейсу Wi-Fi, из-за большей скорости передачи данных, и разработаны основные элементы для ее реализации.

Для передачи параметров сигнала с компьютера на устройство, последнее должно иметь соответствующие аппаратные модули. Для выбора технического решения были протестированы следующие решения: плата Arduino с Wi-Fi модулем ESP8266 ESP-01, плата NodeMCU с микроконтроллером Esp8266, плата Super Mini с ядром ESP32-C3 [2].

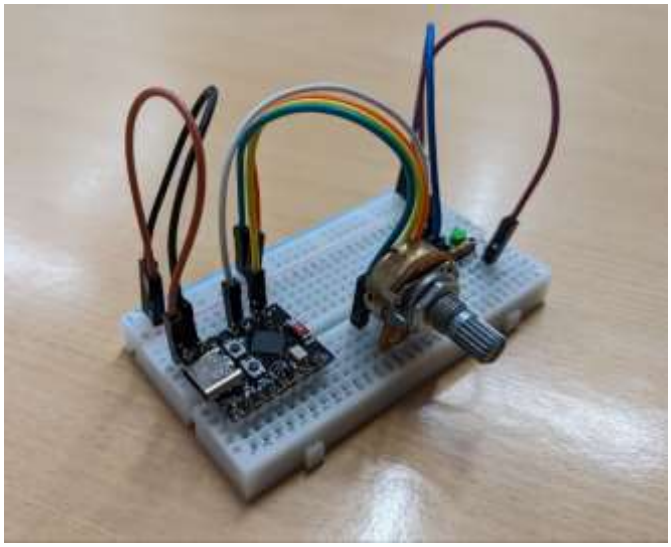


Рис. 1. Плата Super Mini ESP32-C3 с подключенным переменным резистором и тактовой кнопкой

Выбор был сделан в пользу платы Super Mini с ядром ESP32-C3 в силу ее равенства по характеристикам с перечисленными вариантами, но меньшим размерам (рис. 1). Такая плата соответствует всем требованиям: имеет контроллер ШИМ с 6 каналами и разрешением до 14 бит, встроенный Wi-Fi модуля, размер, скорость и стабильность передачи данных. Более того, ESP32-C3 Super Mini обладает низким энергопотреблением, что делает эту плату отличным вариантом для носимого устройства.

Для реализации системы было разработано приложение с серверной частью для получения и обработки данных, отправленных с микроконтроллера и реализована клиентская часть на базе микроконтроллера ESP32-C3.

Приложение должно обеспечивать удобный интерфейс для настройки и управления параметрами. Были рассмотрены и протестированы следующие фреймворки: Boost, WxWidgets, Qt. Выбор был сделан в пользу Qt в силу кроссплатформенности и обширного функционала. Qt – это библиотека на C++, которая позволяет разрабатывать приложения для различных операционных систем [3].

Интерфейс должен включать в себя: поля для ввода параметров (частоты, скважности, времени воздействия) для трёх СИД, кнопки для запуска и остановки формирования сигнала, журнал событий, список для выбора текущего устройства. Так, разработанный интерфейс был готов к подключению функциональной части (рис. 2).

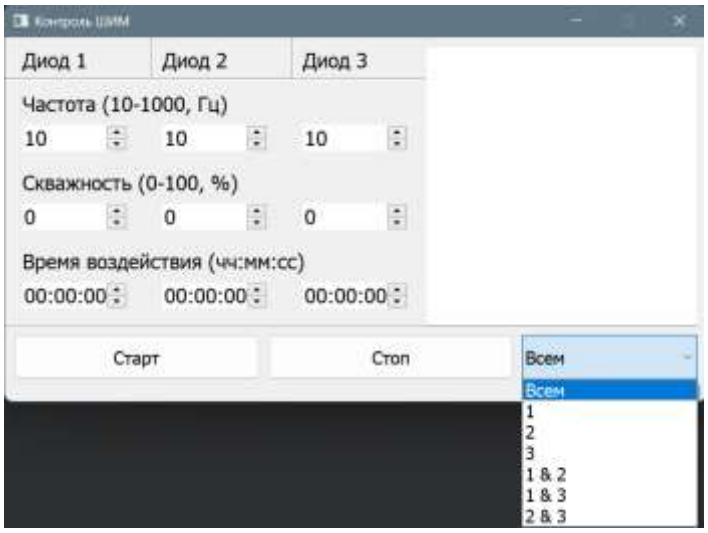


Рис. 2. Интерфейс приложения управления контроллером СИД для модуля ФБМ

Для реализации отправки сообщения контроллеру была разработана серверная часть и использованы сигналы и слоты – ключевой механизм в Qt. После каждой отправки сигнала в журнале событий выводится сообщение с выбранными параметрами сигнала и диодом.

Для единообразия и эффективного считывания данных, полученных от приложения контроллером, был разработан протокол отправки сигнала. Таким образом, контроллер всегда получает одинаковое количество параметров на вход, что упрощает работу программы. Таким образом, было разработано приложение для управления контроллером СИД модуля ФБМ (рис. 2), проверена его работоспособность. После реализации приложения, был выполнен следующий этап работы – разработка клиентской части. Программное обеспечение микроконтроллера должно принимать на вход данные от ПК по Wi-Fi, обрабатывать их и управлять работой СИД.

Таким образом, был реализован механизм, в котором сообщение в специальном формате считывается контроллером по Wi-Fi, разбивается на 3 канала и запускает работу СИД. Далее, пока время воздействия не истечет, или от приложения не будет получен сигнал остановки, сигнал(ы) будет воспроизводиться непрерывно (рис. 3).

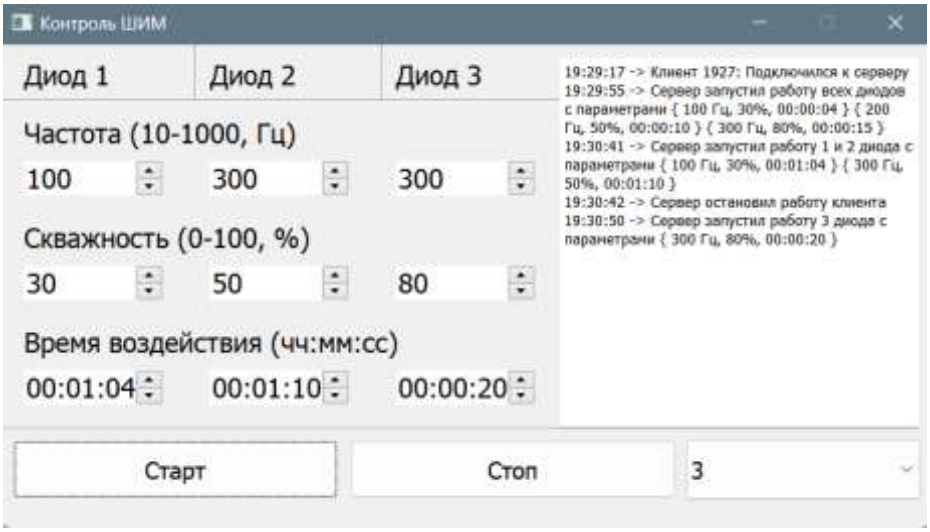


Рис. 3. Итоговый вариант приложения управления контроллером СИД для модуля ФБМ

После того как была разработана основная функция контроллера модуля ФБМ, важно иметь возможность наглядно отследить изменения в работе мозга до и после стимуляции. Для решения этой задачи необходимо приложение, позволяющее производить обработку и управление сигналами ЭЭГ, а также их визуализацию.

Первым этапом было изучение возможности обработки файлов формата ЭЭГ (.edf). В результате поиска решения была найдена библиотека с открытым исходным кодом EDFlib, которая работает на C/C++ [4]. Она предоставляет набор инструментов, которые позволяют упростить механизм получения данных и параметров файлов ЭЭГ.

Так с помощью имеющихся файлов ЭЭГ, была протестирована библиотека и был произведен первый вывод сигнала в цифровом виде. Получившиеся значение были сравнены с популярным приложением для мониторинга ЭЭГ EDFBrowser.

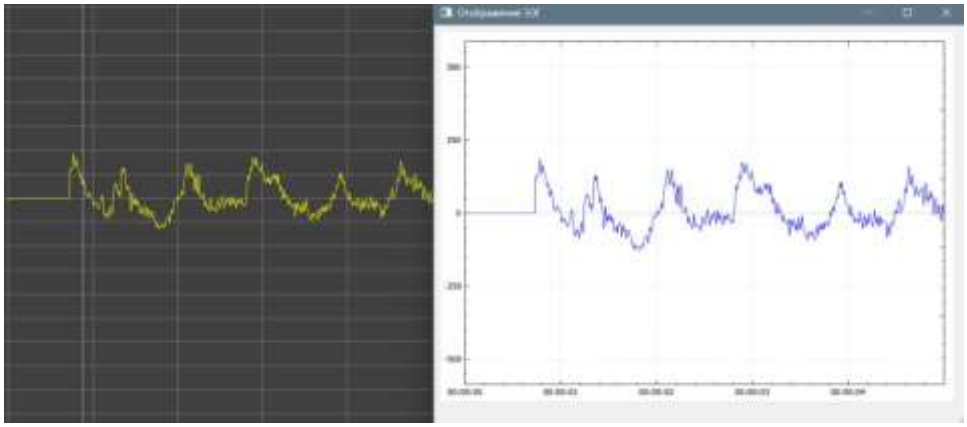


Рис. 4. Вывод значений сигнала в графическом виде. Слева – пример в приложении EDFBrowser, справа – полученный график

Следующим этапом была визуализация сигнала. Для вывода данных в графическом виде можно воспользоваться специальной технологией, которая могла бы решить эту задачу. В качестве такого инструмента была выбрана библиотека QCustomPlot, разработанная для использования в приложениях на C++ с графическим интерфейсом Qt [5]. Таким образом, после подготовки виджета и настройки осей был реализован алгоритм чтения файла и вывода сигнала на экран (рис. 4).

После успешного вывода графика на экран необходимо создать управление, чтобы передвигать камеру по графику. Был разработан интерфейс, и с помощью слотов и сигналов были подключены действия к созданным кнопкам и к ползунку, а также реализованы функции для перемещения и запуск графика в реальном времени.

Следующим этапом была реализация возможности выбора необходимых сигналов и их вывод на экран. Для этого было решено создать список всех доступных сигналов, список выбранных сигналов и кнопки для управления. Был создан механизм добавления нескольких сигналов одновременно, был рассчитан и выставлен промежуток между ними.

Чтобы каждый сигнал был хорошо различим, был реализован механизм выбора контрастного цвета для каждого графика и выбран контрастный фон. Таким образом, ПО для получения данных от устройства ЭЭГ было разработано (рис. 5).

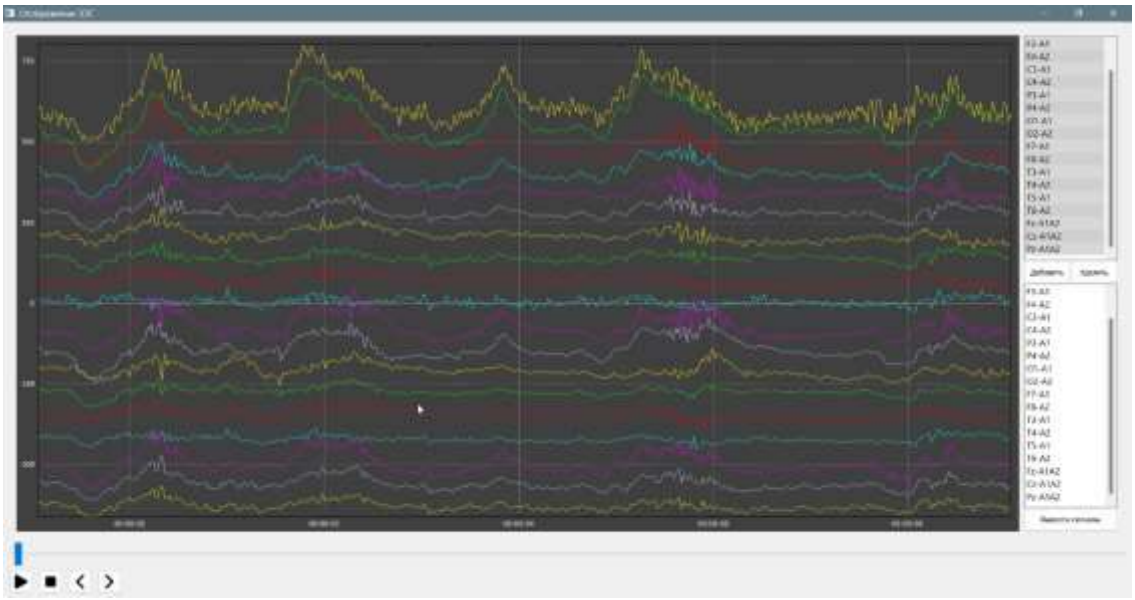


Рис. 5. Итоговый вариант ПО для получения данных от устройства ЭЭГ

Таким образом, в ходе работы были разработаны: контроллер модуля ФБМ, программное обеспечения для мониторинга данных ЭЭГ. Так, помощью разработанной системы специалист сможет стимулировать определенные участки мозга с помощью модуля ФБМ и отслеживать изменения в работе мозга, используя программное обеспечение.

Список литературы

1. Олифер В., Олифер Н. Компьютерные сети. Принципы, технологии, протоколы: Юбилейное издание. – СПб.: Питер, 2021. – 1008 с.
2. ESP32-C3 Series Datasheet. [Электронный ресурс]: URL: https://www.espressif.com/sites/default/files/documentation/esp32-c3_datasheet_en.pdf (дата обращения: 10.10.2024)
3. Qt 5.15. [Электронный ресурс]: URL: https://doc.qt.io/qt-5/ (дата обращения: 18.10.2024)
4. EDFlib for C/C++. [Электронный ресурс]: URL: https://www.teuniz.net/edflib/ (дата обращения: 25.10.2024)
5. QCustomPlot. [Электронный ресурс]: URL: https://www.teuniz.net/edflib/ (дата обращения: 30.10.2024)

АВТОМАТИЗАЦИЯ ETL-ПРОЦЕССОВ С ИСПОЛЬЗОВАНИЕМ APACHE AIRFLOW

Котов И. И., Куприянова А. М.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В статье описывается применение Apache Airflow для автоматизации ETL-процессов в современных проектах. Предложен метод проектирования Directed Acyclic Graphs (DAG) для оркестрации задач, интеграции с разнородными источниками данных (СУБД, API) и повышения надежности пайплайнов.

Ключевые слова: Apache Airflow, ETL, автоматизация, DAG (Directed Acyclic Graph), оркестрация данных, интеграция API, масштабируемость, pipeline.

Введение

С ростом объемов данных и усложнением бизнес-логики ручное управление ETL-процессами становится неэффективным. Традиционные подходы требуют значительных временных затрат на мониторинг, обработку ошибок и интеграцию источников. Организации нуждаются в инструментах, которые обеспечивают:

1. Автоматизацию задач извлечения, трансформации и загрузки данных
2. Гибкое управление зависимостями между этапами обработки
3. Масштабируемость под растущие нагрузки
4. Визуализацию выполнения пайплайнов

Apache Airflow, основанный на концепции DAG, решает эти задачи, позволяя описывать все рабочие процессы кодом на Python [1, 2].

Преимущества Apache Airflow

1. DAG как основа ETL. Задачи объединяются в графы с явными зависимостями. Например, задача извлечения данных из API выполняется перед их валидацией и загрузкой в БД [3, 4].

2. Гибкость через Python-код. Задачи описываются программно, что позволяет реализовать сложную бизнес-логику, использовать различные библиотеки и настраивать триггеры.

3. Широкая экосистема интеграций. Поддержка СУБД (PostgreSQL, MySQL), облачных хранилищ (AWS S3, Google Cloud) и BI-инструментов (Tableau).

4. Масштабируемость. Возможность запуска в распределенных средах (Kubernetes, Celery) и настройки большого числа параметров среды.

5. Визуализация и мониторинг. Веб-интерфейс предоставляет графическое представление DAG, логи выполнения и метрики.

Разработанный ETL-пайплайн

В рамках проекта реализованы два ключевых DAG-а (рисунок 1), демонстрирующие гибкость Airflow в оркестрации данных. Важно отметить, что предложенное решение позволяет снять нагрузку, например, с аналитиков компании за счет автоматизации процессов сбора и обработки данных. При этом добавляется возможность контроля за процессами с помощью DAG-а мониторинга.

1. DAG tgsell_loader

Автоматизирует сбор данных о новых пользователях и магазинах через открытое API стороннего сервиса tgsell (был взят для примера), а также загружает их в PostgreSQL - демонстрация на рисунке 2.

Механизм работы:

а) Извлечение (Extract): динамические задачи, созданные кодом через цикл, запрашивают данные из API

б) Трансформация (Transform): валидация данных с использованием Pydantic-моделей для проверки структуры и качества

в) Загрузка (Load): сохранение данных в PostgreSQL

2. DAG dags_restarter

Повышает отказоустойчивость и эффективность системы за счет автоматического перезапуска неудачных задач, а также за счет логирования информации о соответствующих

процессах. На рисунке 3 есть розовые ячейки - так отмечаются выполнения, в процессе которых был совершен перезапуск какого-либо из DAG-ов, также видно, что в логах отображается подробная информация о последних запусках соответствующего DAG. Для добавления очередного DAG-а на мониторинг достаточно просто вписать его название в конфиг в веб-интерфейсе Airflow (рис. 4).

Механизм работы:

а) Мониторинг статусов DAG

б) Автоматический перезапуск задач при обнаружении ошибок

в) Понятное логирование всей информации в веб-интерфейс Airflow

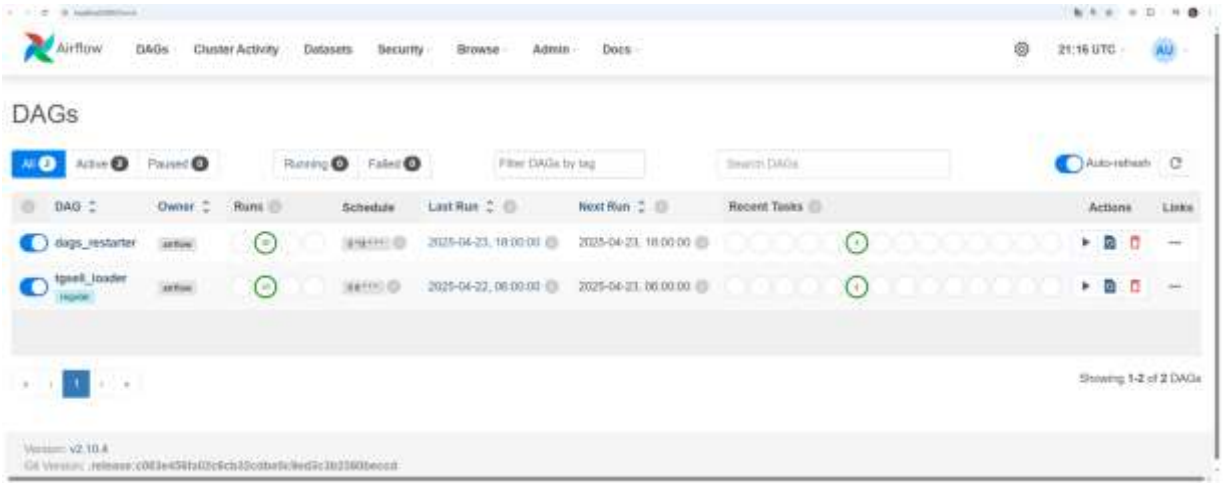


Рис. 1. DAG-и - главная страница UI Airflow

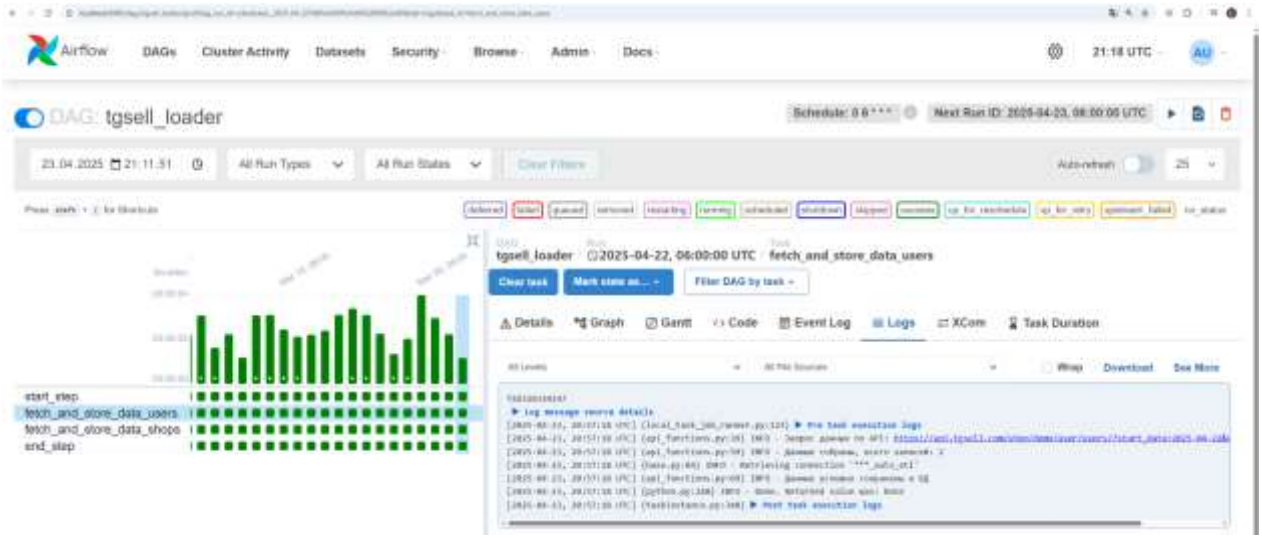


Рис. 2. DAG tgsell_loader

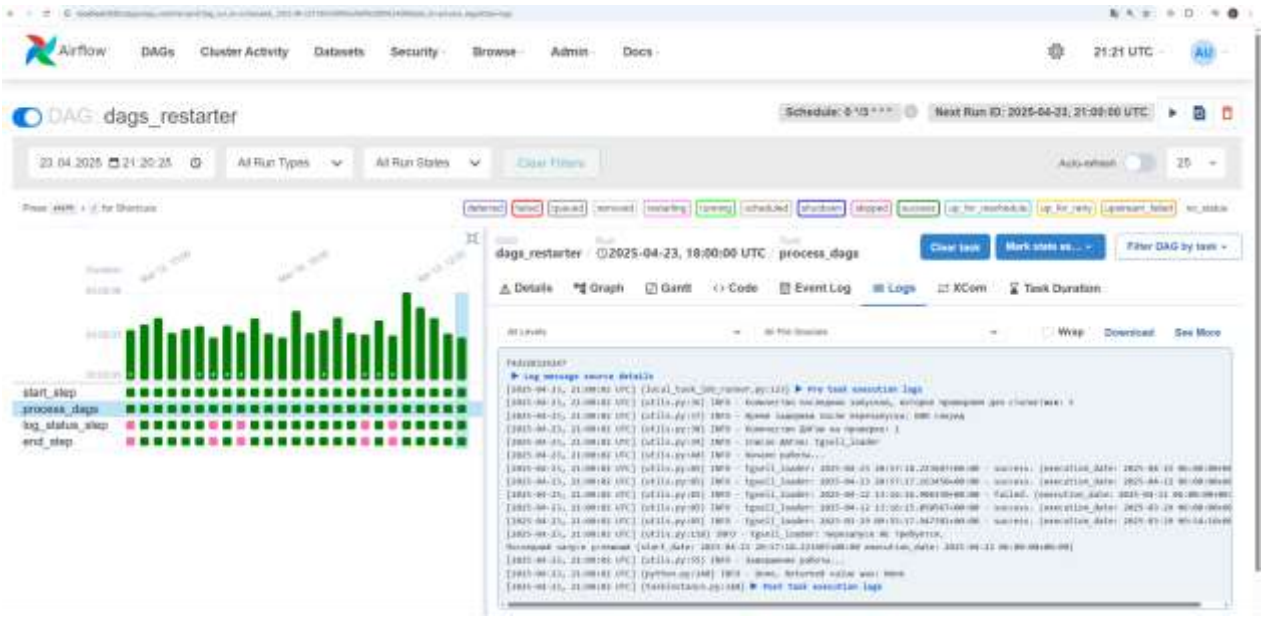


Рис. 3. DAG dags_restarter

Рис. 4. Настройка переменных окружения

Результаты внедрения

1. Сокращение времени обработки. Автоматизация ETL позволила почти полностью исключить ручные операции.

2. Повышение точности данных. Валидация через Pydantic исключила расхождения при сборе и обработке данных.

3. Улучшение стабильности системы. Ежедневный запуск по расписанию соответствующих DAG-ов позволяет сохранять консистентность данных в корпоративном хранилище.

4. Гибкость конфигурации. Использование Airflow Variables и Connections обеспечило безопасное хранение настроек с возможностью разграничения доступов.

Заключение

Apache Airflow доказал свою эффективность в автоматизации ETL-процессов. Реализованные решения, такие как интеграция с API, качественная валидация данных и автоматический перезапуск задач по заданным условиям, не только оптимизируют текущие операции, но и обеспечивают основу для масштабирования, закладывая фундамент для дальнейших разработок.

Список литературы

1. Apache Airflow Documentation. Airflow Components Overview [Электронный ресурс] // Apache Software Foundation. URL: https://airflow.apache.org/docs/apache-airflow/stable/core-concepts/overview.html#airflow-components (дата обращения: 28.02.2025).
2. Airbnb Engineering: Scaling ETL with Airflow [Электронный ресурс] // Airbnb Engineering & Data Science. URL: https://medium.com/airbnb-engineering/data-airflow-a-workflow-management-platform-46318b977fd8 (дата обращения: 27.02.2025).
3. Kimball R., Ross M. The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data [Электронный ресурс] // Wiley, 2004. URL: https://www.r-4.org/files/books/computers/databases/toolkits/Ralph_Kimball_Joe_Caserta-The_DWH_ETL_Toolkit-EN.pdf (дата обращения: 27.02.2025).

АВТОМАТИЗАЦИЯ А/Б-ТЕСТИРОВАНИЯ

Куприянова А. М., Котов И. И., Куприянов Н. М.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Данная статья посвящена описанию ручного и автоматизированного подходов к А/Б-тестированию в контексте цифровой трансформации бизнеса. Рассматриваются преимущества и недостатки ручного и автоматизированного тестирования. Акцент сделан на автоматизации процесса, которая ускоряет тестирование и повышает точность результатов. Предлагается гибкое решение по автоматизации, позволяющее проводить полноценное А/Б-тестирование.

Ключевые слова: А/Б-тестирование, автоматизация процессов, ручное тестирование, no-code платформы, цифровая трансформация, статистическая значимость, машинное обучение, e-commerce, data-driven управление, оптимизация конверсии, маркетинговые исследования

Введение

Цифровая трансформация в настоящее время один из главных трендов в бизнес компаниях, которая проводится для повышения эффективности процессов. В условиях высокой конкуренции маркетологам и продакт-менеджерам приходится почти ежедневно проводить тестирование гипотез о влиянии различных экономических метрик на конверсию и прибыль, а также выбирать наиболее оптимальный вариант дизайна сайта или приложения, чтобы привлечь максимальное количество пользователей и т.д. Общепринятым методом для решения этих задач является А/Б-тестирование — метод маркетингового исследования, позволяющий на основе статистики оценить влияние изменения на метрики продукта [1].

Целью данной статьи является описание подходов к А/Б-тестированию, а также демонстрация преимуществ автоматизации данного процесса.

Описание ручного А/Б-тестирования

Ручное тестирование – первое, что выбирает большинство компаний на начальной стадии разработки приложения или сайта. Оно хорошо подходит для краткосрочных и небольших проектов, которые могут быть ограничены в бюджете, а также оно все еще является лучшим вариантом для интуитивного или исследовательского тестирования, когда точность результатов во многом зависит от реального взаимодействия пользователя с продуктом.

Однако, несмотря на простоту и доступность метода, у него есть и минусы. Он может занимать много времени, не справляется с большими объемами данных, не поддерживает

мультивариантное тестирование, данные не могут автоматически сихронизироваться с CRM и BI-инструментами. Также большим недостатком является субъективность анализа: по мнению Томи Местера часто маркетолог выбирает вариант B, потому что «он красивее», а не из-за статистической значимости p-value [2].

Описание автоматизированного А/Б-тестирования

Автоматизированное А/Б-тестирование чаще всего используют более крупные компании, так как переход от трудоемких ручных процессов к управляемым алгоритмам особенно удобен при росте объема данных и потребности в быстрых итерациях.

Автоматизацию можно реализовать собственными решениями, написанными средствами Python, а также при помощи специальных no-code платформ, не требующих навыков программирования.

Автоматизация позволяет проводить тестирование за более короткое время и повышать точность результатов за счет моделей машинного обучения и более продвинутых методов статистического анализа. Это в своем исследовании доказал Дэниэл Л. Асфау: он смоделировал два тестовых случая, помогающие сравнить эффективность, анализируя лучшее время теста, достигнутое как ручным, так и автоматизированным способом.[3].

Таблица 1

Результаты сравнения эффективности подходов к А/Б-тестированию

Test case	Manual best test time	Automation test time
Test case 1	300 seconds	20 seconds
Test case 2	459 seconds	30 seconds

Из получившейся таблицы видно, что время выполнения обоих тестовых кейсов было в несколько раз меньше при автоматизации, чем при ручном подходе.

Также Асфау исследовал точность результатов тестирования и установил, что с ростом уровня компетенций тестировщиков точность возрастала, но само наличие компетенции не гарантировало высокую точность в отличие от автоматизации, при которой ожидаемый результат был достигнут в 100% случаев.

К недостаткам решений по автоматизации, относятся более высокие стартовые затраты на реализацию и потребность в квалифицированных разработчиках, однако это окупается за счет кастомизации автоматизированного решения.

При использовании же no-code платформ, которые подходят большинству компаний любого масштаба, таких проблем не возникает. Они экономят время и ресурсы и не требуют навыков программирования. Несмотря на меньшую гибкость в сравнении с кастомными решениями, no-code платформы — это идеальный старт для автоматизации A/B-тестов. Они дают возможность даже небольшим командам конкурировать с крупными игроками.

На сегодняшний день существует множество no-code платформ, самыми популярными и доступными для российских пользователей являются: LPgenerator, Roistat и Mindbox и Callibri. [4]

Описание предлагаемого решения для автоматизации тестирования

В рамках проекта, был создан инструмент для автоматизации, реализующий полный цикл тестирования: от сбора данных и их очистки до статистического анализа и интерпретации результатов.

Основные этапы проекта:

1. Генерация синтетических данных с эффектами для групп A/B
2. Обработка данных:
   - Удаление пропусков
   - Очистка от дубликатов (полных, частичных, семантических)
   - Удаление выбросов
3. Статистический анализ:
   - Проверка баланса групп
   - Z-тесты и T-тесты
   - Расчет MDE (минимального детектируемого эффекта)
4. Интеграция с MinIO и PostgreSQL для хранения файлов
5. REST API на FastAPI

Архитектура системы реализована как модульный конвейер данных, включающий следующие компоненты:

Сбор данных → Обработка → Статистический анализ → Интеграция → Визуализация

В скриншотах, представленных ниже продемонстрированы ключевые результаты работы программы:

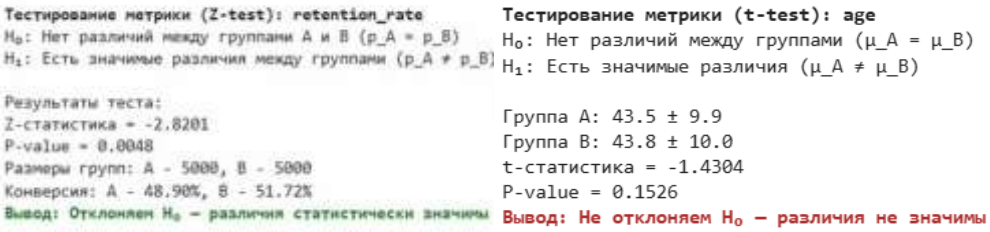


Рис. 1. Результаты проведения t-теста и Z-теста

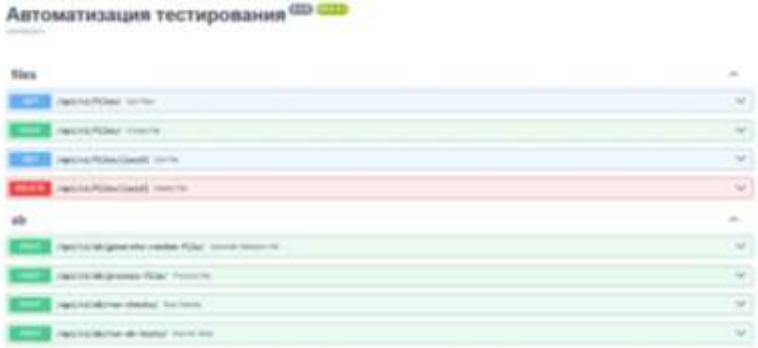


Рис. 2. API

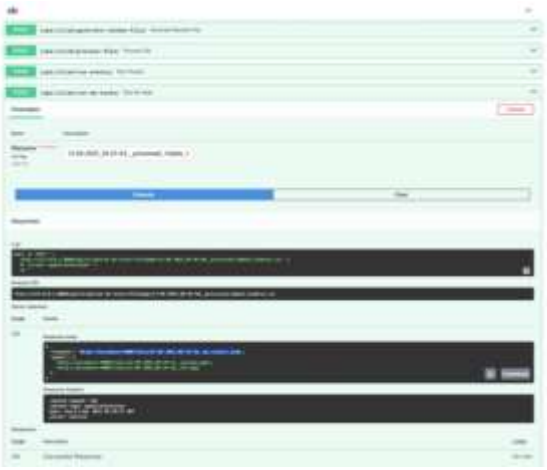


Рис.3. Результаты тестирования после методов обработки

По итогам исследования можно сформировать сравнительную таблицу для описанных выше подходов, которая наглядно демонстрирует преимущество не только автоматизации в целом, но также и преимущество no-code платформ перед ручным тестированием. Дополнительные сведения о сравнении можно почерпнуть из [3].

Таблица 2

Сравнение методов проведения А/Б-тестирования

Критерий	**Ручное А/Б-тестирование**	**Автоматизированное А/Б-тестирование**	
		Кастомные решения	**No-code платформы**
Время выполнения	Дни/недели	Часы/дни	Часы/дни
Стартовые затраты	Низкие	Высокие	Умеренные или низкие
Объем данных	Малый	Большой	Средний и малый
Технические требования	Минимальные	Высокие	Нулевые
Интеграции	Ручной перенос данных	Автоматическая синхронизация через API	Готовые коннекторы, CRM, BI-системы
Скорость анализа	Медленная	Высокая	Высокая
Точность результатов	Зависит от навыков тестировщика	Максимальная	Высокая
Функционал тестирования	Широкий	Широкий	Ограниченный
Тип бизнеса	Малый	Крупный и средний	Средний и малый

Заключение

На основании анализа и полученных результатов можно сделать следующие выводы:

1. Выбор подхода для А/Б-тестирования определяется в первую очередь масштабом бизнеса и целями компании, а также ожидаемой точностью результатов тестирования. Однако тренд на автоматизацию становится неизбежным этапом эволюции компаний, стремящихся к data-driven управлению и устойчивому росту в цифровой эпохе.
2. No-code платформы способны заменить ручной подход, но обладают ограниченным функционалом и не подходят для большого объема данных.
3. Создание автоматизированного кастомного решения требует более высоких затрат на старте, но позволяет существенно экономить время тестирования, повысить точность результатов и анализировать большие объемы данных.

4. Разработанная программа является качественным примером и прототипом кастомного решения, удовлетворяющим базовым требованиям автоматизации тестирования (сбор данных, обработка, статистический анализ, интеграция, визуализация).

Список литературы

1. A/B-тестирование [Электронный ресурс]. // Википедия // URL: https://ru.wikipedia.org/wiki/A/B-%D1%82%D0%B5%D1%81%D1%82%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5 (дата обращения: 17.04.2025).
2. Statistical Significance in A/B testing (and How People Misinterpret Probability) [Электронный ресурс]/ Tomi Mester // data36 // URL: https://data36.com/statistical-significance-in-ab-testing/ (дата обращения: 17.04.2025).
3. Benefits of Automated Testing Over Manual Testing [Электронный ресурс]/ Daniel L Asfaw //Academia // URL: https://www.academia.edu/26811709/Benefits_of_Automated_Testing_Over_Manual_Testing (дата обращения: 17.04.2025).
4. A/Б-тестирование: 6 ТОП инструментов для проведения [Электронный ресурс]/ Никита Соловов // vc.ru // URL: https://vc.ru/niksolovov/1156951-a-b-testirovanie-6-top-instrumentov-dlya-provedeniya-2025 (дата обращения: 18.04.2025).

РАЗРАБОТКА СИСТЕМЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ ТЕМПЕРАТУРОЙ ПОДЛОЖКОДЕРЖАТЕЛЯ НА УСТАНОВКЕ ВУП-11М

Т. И. Мавлявиев, К. М. Моисеев

Московский Государственный Технический Университет им. Н. Э. Баумана

Аннотация. В статье представлена система автоматического управления температурой подложкодержателя вакуумной установки ВУП-11М, предназначенной для нанесения тонкоплёночных покрытий. Описаны принципы реализации ПИД-регулятора на базе микроконтроллера ESP32. Передача данных между модулями организована по беспроводному протоколу UDP. Приведена электрическая схема подключения компонентов. Показано, как разработанная система может быть внедрена параллельно действующей системой автоматического и ручного управления без вмешательства в её структуру.

Ключевые слова: система автоматического управления, беспроводная передача данных, ПИД-регулятор, микрокомпьютер, микроконтроллер, вакуумная технологическая установка, нанесение тонких плёнок

Введение

Температура подложки в процессе формирования тонкоплёночных покрытий оказывает критическое влияние на морфологию, структуру и адгезию осаждаемых плёнок, определяя тем самым их функциональные свойства. На вакуумной установке магнетронного распыления ВУП-11М, используемой в лаборатории «Вакуумные технологии наноинженерии» МГТУ им. Н.Э. Баумана, нагрев подложки осуществляется резистивным нагревателем, расположенным в подложкодержателе [1].

До внедрения автоматической системы регулирования управление температурой осуществлялось вручную. Предварительно проводились дополнительные эксперименты, в ходе которых оценивались параметры работы установки и подбиралась мощность на блоке питания нагревателя для достижения необходимой температуры. Такой подход отличался большими затратами времени и не очень высокой точностью.

Для повышения стабильности технологического процесса и снижения влияния человеческого фактора было принято решение о разработке системы автоматического управления температурой подложкодержателя, основанной на ПИД-регулировании.

Целью настоящей работы является разработка простой и надёжной системы автоматического регулирования температуры, способной работать параллельно с существующими блоками управления без вмешательства в их структуру.

Объект управления

Объектом управления в разрабатываемой системе является подложкодержатель, расположенный внутри рабочей камеры установки. Его конструкция определяет как тепловые характеристики процесса, так и способ размещения датчиков и элементов управления температурой.

Рис. 1. Подложкодержатель с нагревом

Подложкодержатель изготовлен из нержавеющей стали и состоит из нескольких частей: верхней – рефлектора, резистивного нагревателя и нижней пластины.

Верхняя часть представляет собой рефлектор, расположенный непосредственно над нагревательным элементом. Его основная функция — отражение теплового излучения от резистивного нагревателя обратно к размещённым ниже подложкам. Через отверстия в рефлекторе проходят провода для питания нагревателя и съёма информации со встроенной в нагреватель термопары.

Нижняя часть — это прямоугольная пластина, к которой крепятся подложки. При проведении процессов, в которых требуется нагрев для надёжного крепления используются специальные металлические лапки, для которых в пластине предусмотрены четыре монтажных отверстия. Вместо одной из лапок устанавливается термопара типа K с изолированным спаем, обеспечивающая измерение температуры непосредственно в области размещения подложек.

Таким образом, конструкция подложкодержателя позволяет контролировать температуру как в зоне нагревателя, так и в зоне технологической поверхности.

Аппаратная и программная реализация

Разрабатываемая система управления включает в себя микроконтроллер ESP32, усиливающий модуль MAX31855 и одноплатный компьютер Raspberry Pi (RPi).

Термопара типа K подключается к модулю MAX31855, который формирует цифровой сигнал, считываемый микроконтроллером ESP32. Контроллер с заданным интервалом отправляет пакет с текущим значением температуры и временной меткой по протоколу UDP на RPi.

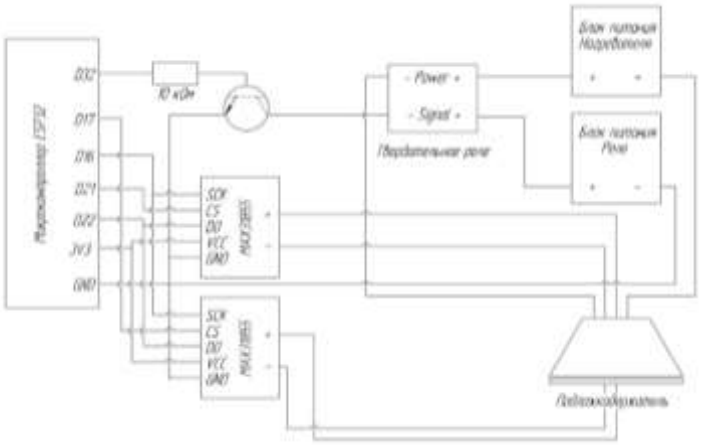


Рис. 2. Типовая схема системы нагрева и контроля температуры

Поскольку регулирование осуществляется непосредственно на ESP32, Raspberry Pi играет вспомогательную роль: он принимает данные, визуализирует их и может отправлять команды для изменения целевой температуры или включения/отключения регулятора.

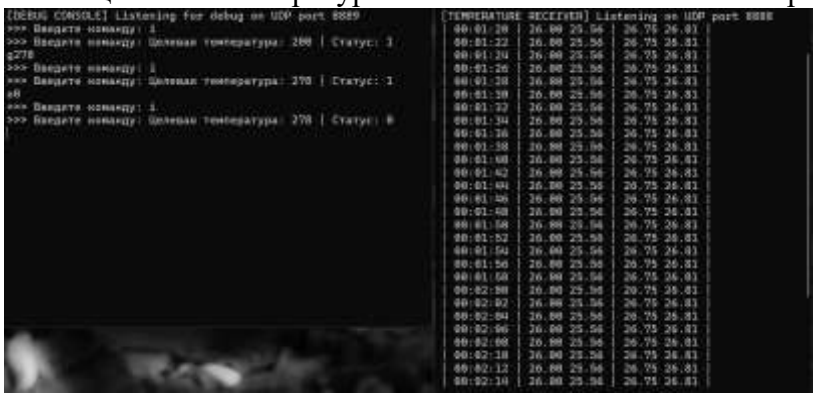


Рис. 3. Интерфейс управления на Raspberry Pi

Работа и настройка ПИД-регулятора

В разработанной системе автоматического регулирования температуры подложкодержателя применяется ПИД-регулятор (пропорционально-интегрально-дифференциальный), реализованный программно на микроконтроллере ESP32 [2].

Алгоритм ПИД-регулятора основан на классическом выражении:

$$u(t) = P + I + D = K_p e(t) + K_i \int_0^t e(\tau)\,d\tau + K_d \frac{de}{dt},$$

где $u(t)$ – выходной сигнал регулятора;

$e(t)$ – ошибка регулирования;

K_p, K_i, K_d – коэффициенты усиления пропорциональной, интегрирующей и дифференцирующей составляющих регулятора соответственно [3, 4].

Измеренная температура поступает с термопары типа K через усилитель, а управляющее воздействие подаётся в виде ШИМ-сигнала на твердотельное реле, управляющее нагревателем. Скважность сигнала определяет подводимую мощность и, соответственно, температуру подложкодержателя. Период ШИМ-сигнала также рассчитывается. Это обеспечивает более высокую точность поддержания температуры при

области близкой к целевой была выше, при этом учитывается максимальная частота работы реле в 50 Гц.

Настройка коэффициентов ПИД-регулятора выполнялась с использованием библиотеки открытой sTune, которая реализует полуэмпирический подход к автоматической настройке [5]. Из прошлого опыта работы с установкой были определены постоянная времени и некоторые другие необходимые для расчётов параметры. Контроллер вычислял коэффициенты по заранее заданным данным для определённого метода настройки (Зиглера-Николя, C-H-R и другие). Настройка проводилась по данным с термопары, расположенной на пластине для подложек [6].

Тестирование и результаты

Работоспособность системы была проверена в процессе нагрева подложкодержателя на установке ВУП-11М. Тестирование проводилось для коэффициентов, вычисленных разными методами. Коэффициенты, полученные методом C-H-R после небольшого редактирования, показали лучшие результаты.

На этапе запуска системы проводилось достижение заданной температуры в диапазоне 500 °C. ПИД-регулятор обеспечивал выход на уставку с незначительной статистической ошибкой и допустимым временем установления.

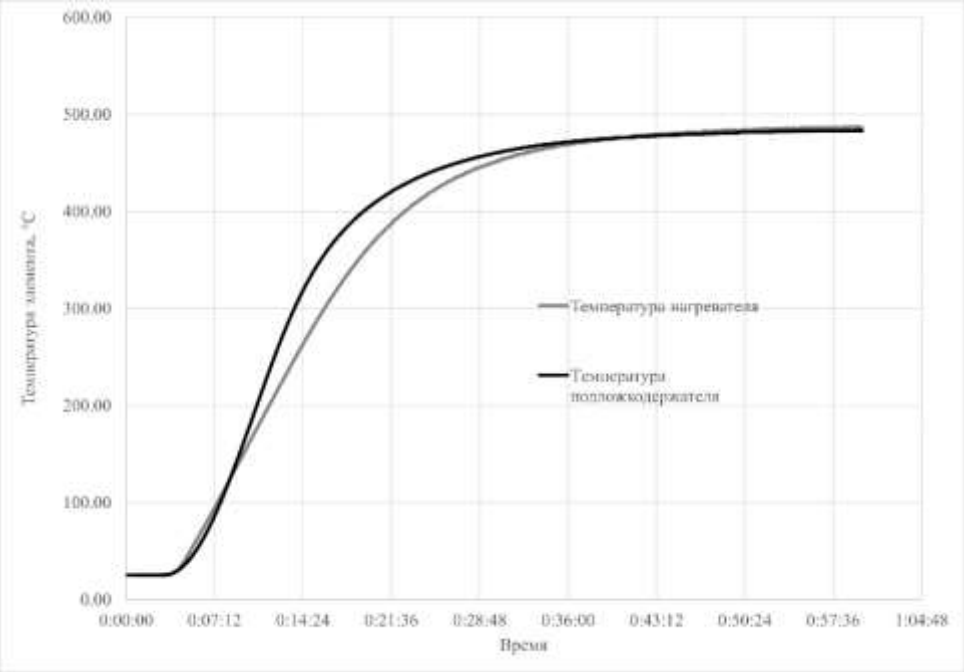


Рис. 4. Графики изменения температуры элементов

Полученные температурные кривые демонстрируют типичную динамику работы ПИД-регулятора с плавным выходом на уставку и последующей стабилизацией температуры.

Беспроводная передача данных по протоколу UDP оказалась надёжной: за время тестов потери пакетов на время большее десяти секунд не наблюдалось, что обусловлено малой загруженностью локальной сети.

Заключение

В результате проделанной работы разработана и протестирована система автоматического управления температурой подложкодержателя на установке ВУП-11М, основанная на ПИД-регулировании с реализацией на микроконтроллере ESP32.

Система обеспечивает стабильный температурный режим и может быть внедрена параллельно существующей САУ без изменений в её аппаратной части, что делает её удобной для модернизации лабораторного оборудования.

Список литературы

1. Моисеев, К.М. Многопозиционные вакуумные технологические установки // Наноинженерия. 2014. № 8. С. 18–21.
2. Bennett S. A brief history of automatic control //IEEE Control Systems Magazine. – 2002. – Т. 16. – №. 3. – С. 17-25.
3. Ang K. H., Chong G., Li Y. PID control system analysis, design, and technology //IEEE transactions on control systems technology. – 2005. – Т. 13. – №. 4. – С. 559-576.
4. Sheel S., Gupta O. New techniques of PID controller tuning of a DC motor-development of a toolbox //MIT International Journal of Electrical and Instrumentation Engineering. – 2012. – Т. 2. – №. 2. – С. 65-69.
5. sTune [Электронный ресурс]. – Режим доступа: https://github.com/Dlloydev/sTune/tree/main - Заглавие с экрана. – (Дата обращения: 10.05.2025).
6. Basu A., Mohanty S., Sharma R. Tuning of FOPID controller for meliorating the performance of the heating furnace using conventional tuning and optimization technique //International journal of electronics engineering research. – 2017. – Т. 9. – №. 1. – С. 69-85.

ОРГАНИЗАЦИЯ ВЗАИМОДЕЙСТВИЯ С ДАТЧИКАМИ МОНИТОРИНГА ДОБЫВАЕМОЙ СМЕСИ НЕФТЕДОБЫВАЮЩИХ СКВАЖИН

МАКАРОВ М.В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В статье рассматривается способ взаимодействия между персональным компьютером и регистратором данных, установленным на нефтепроводе с целью получения вибрационных, акустических и температурных данных. Взаимодействие производится при помощи интерфейса RS-232 и протокола UDP. Полученные программы тестируются при помощи разработанного эмулятора регистратора данных.

Ключевые слова: регистратор данных, RS-232, UDP, OSI

Введение

В современном мире существует большое количество сфер, требующих сбора значительного количества информации. Нефтяная отрасль не является исключением. Анализ данных, поступающих с датчиков давления и температуры необходим для определения состояния скважины, определения потока, проходящего по трубе, предупреждения неисправностей и реагирования на чрезвычайные ситуации [1, 2, 3, 4]. Для взаимодействия с такими регистраторами данных необходима разработка специального программного обеспечения, предоставляющего удобный программный интерфейс, позволяющий упростить управление регистратором, сохранение большого объёма данных и их интерпретацию.

Цель исследования — разработать систему взаимодействия с регистратором данных (микроконтроллером). Для реализации поставленной цели необходимо выполнить следующие задачи.

1. Реализовать механизм передачи команд между персональным компьютером и микроконтроллером.
2. Реализовать механизм передачи данных между микроконтроллером и приемником данных.
3. Реализовать пользовательский интерфейс для управления процессом приёма и передачи данных.
4. Провести тестирование разработанной системы с помощью модельного эксперимента.

Методы реализации

Для реализации передачи команд к микроконтроллеру требуется выбрать интерфейс физического уровня модели OSI (The Open Systems Interconnection model) [5]. В исследовании были рассмотрены интерфейсы RS-232 [6], RS-485 и USB. Кроме этого, для передачи данных между регистратором и приёмником необходимо выбрать протокол транспортного уровня модели OSI. В исследовании были рассмотрены протоколы UDP (User Datagram Protool) [7] и TCP (Transmission Control Protocol).

Сравнение интерфейсов физического уровня приведено в таблице 1, сравнение протоколов транспортного уровня приведено в таблице 2.

Таблица 1

Интерфейсы физического уровня модели OSI

Параметр	RS-232	RS-485	USB
Тип связи	Точка-точка	Шина	Хост-устройство
Дальность	До 15 метров	До 1200 метров	До 5 метров
Скорость передачи	До 19.2 кбит/с	До 10 Мбит/с	До 480 Мбит/с (USB 2.0)
Помехоустойчивость	Низкая	Высокая (дифференциальный сигнал)	Высокая (дифференциальный сигнал)
Сложность реализации	Низкая	Средняя	Высокая
Дуплексность	Полный дуплекс	Полный дуплекс при использовании дополнительной логики	Полный дуплекс (USB 2.0)

Таблица 2

Протоколы транспортного уровня модели OSI

Параметр	UDP	TCP
Тип соединения	Без установления соединения	С установлением соединения
Гарантия доставки	Не гарантирует доставку данных получателю	Существует гарантия доставки данных получателю
Повторная передача данных	В случае потери пакета данных, повторная передача не осуществляется	В случае потери пакета данных осуществляется повторная передача данных
Проверка ошибок	Базовая проверка ошибок	Полная проверка ошибок

Контроль потока	Отсутствие контроля потока	Осуществляется контроль потока
Скорость	Высокая	Ниже из-за накладных расходов
Применение	Real-time, потоковое видео	Передача файлов, web, электронная почта

В исследовании было решено использовать интерфейс RS-232 для передачи команд от персонального компьютера к регистратору данных из-за простоты его реализации и простоты поддержки со стороны микроконтроллера.

Для передачи данных от микроконтроллера к приёмнику было решено использовать протокол UDP. Данное решение также обосновано простотой реализации со стороны микроконтроллера. Кроме этого, UDP позволяет значительно упростить взаимодействие с потоковыми данными, генерируемыми регистратором из-за отсутствия необходимости подтверждения передачи данных.

Модель системы регистрацции данных

Со стороны ПК была разработана программа, позволяющая запускать различные сценарии использования регистратора данных (рисунок 2). Программа последовательно пересылает команды через интерфейс RS-232 и ждет ответа от регистратора. В случае положительного ответа, программа продолжает выполнение сценария.

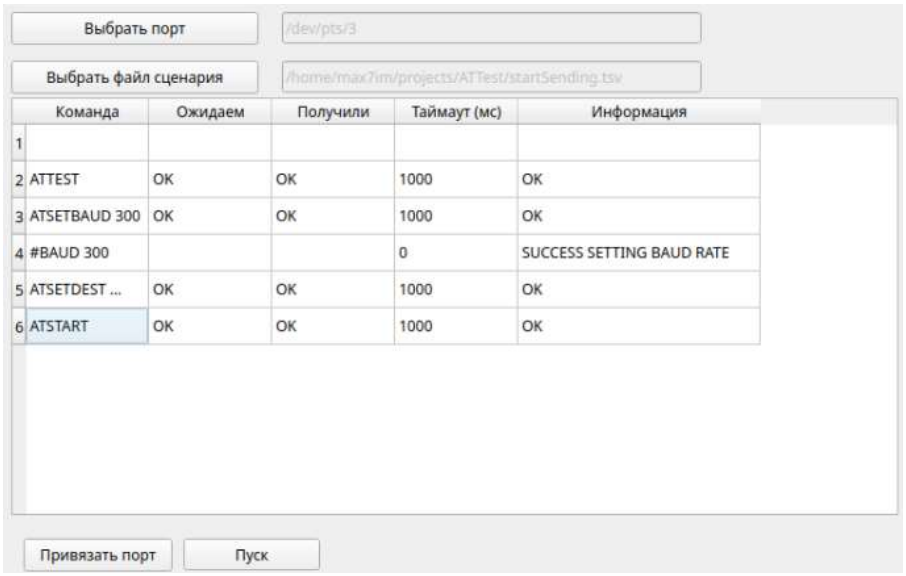


Рис. 2. Программа запуска сценариев

Со стороны приёмника была разработана программа, позволяющая принимать и записывать данные, приходящие на UDP-порт с целью их каталогизации и дальнейшего анализа (рисунок 3). Программа разбивает данные по тому, с какого адреса пришли датаграммы, затем по порту и далее по времени, в которое пакеты были приняты. Принятые данные далее могут быть преобразованы в файл формата wav.

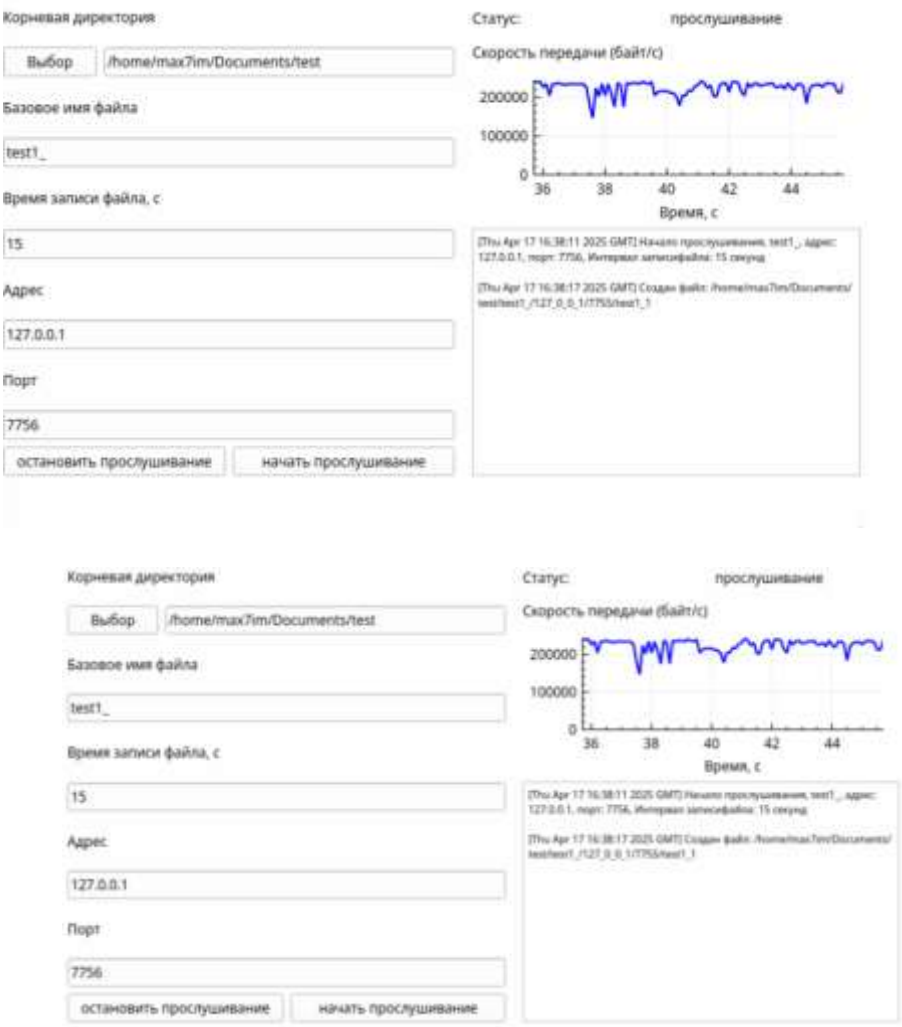


Рис. 3. Программа приёма данных

Для проведения модельного эксперимента также был реализован эмулятор регистратора данных, взаимодействующий с RS-232 и UDP (рисунок 4).

```
max7im@max7im-HP-ProBook-4720s:~/projects/boardEmulator/build$ ./boardEmulator /dev/pts/4
127.0.0.1 7755
Started UDP thread
Started COM thread
running on  "/dev/pts/4"  and  "127.0.0.1" :  7755
com worker started
command:  "ATTEST" , return:  "OK"
command:  "ATSETBAUD 300 " , return:  "OK"
command:  "ATSETDEST 127.0.0.1 7756" , return:  "OK"
set new destination  "127.0.0.1"   7756
udp worker started
command:  "ATSTART" , return:  "OK"
udp worker stopping
udp worker stopped
command:  "ATSTOP" , return:  "OK"
```

Рис. 4. Эмулятор регистратора данных

В ходе проведения модельного эксперимента реализованные модули были соединены в соответствии с рисунком 1. После этого был запущен тестовый сценарий, инициирующий передачу данных с эмулятора на приёмник. По истечении некоторого времени был

активизирован сценарий, останавливающий передачу данных. Переданные данные были преобразованы в файл формата wav (рисунок 5) и проверены на наличие ошибок. В файле присутствуют 11 каналов, среди которых 4 для данных с вибродатчиков, 5 для данных с датчиков температуры и 2 для данных с датчиков давления.

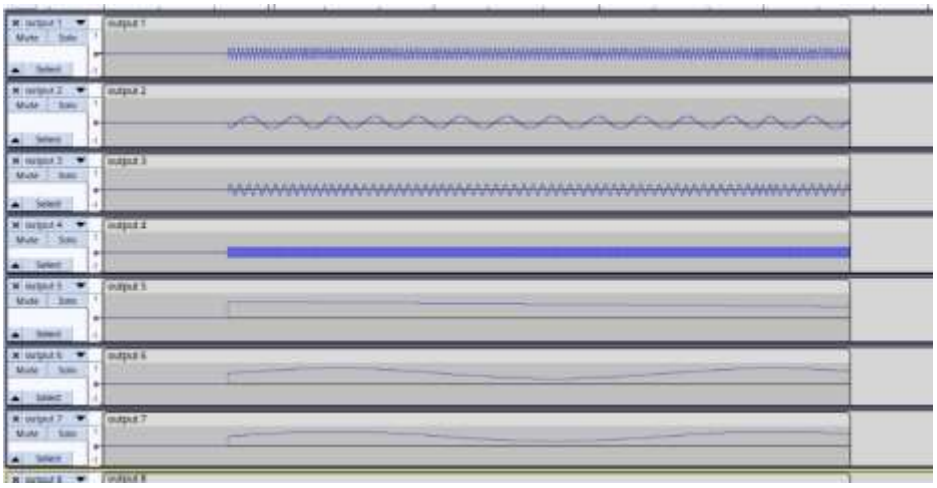


Рис. 5. Результирующий файл в программе audacity

В итоге исследования была создана система взаимодействия с регистратором данных. Система прошла испытание на модельном эксперименте и показала свою работоспособность.

Список литературы

1. Takeda Y. Velocity profile measurement by ultrasound Doppler shift method //International journal of heat and fluid flow. – 1986. – Т. 7. – №. 4. – С. 313-318.
2. Falcone G., Hewitt G., Alimonti C. Multiphase flow metering: principles and applications. – Elsevier, 2009. – Т. 54.
3. Arsalan M. Advances in clamp-on flow measurement techniques //Abu Dhabi International Petroleum Exhibition and Conference. – SPE, 2020. – С. D021S028R004.
4. Venkata S. K., Navada B. R. Estimation of flow rate through analysis of pipe vibration //acta mechanica et automatica. – 2018. – Т. 12. – №. 4. – С. 294-300.
5. Microsoft Corporation. Windows network architecture and the OSI model [Электронный ресурс]. – Microsoft Learn. – 2023. – Режим доступа: https://learn.microsoft.com/en-us/windows-hardware/drivers/network/windows-network-architecture-and-the-osi-model (дата обращения: 01.05.2025).
6. Cami Research Inc. RS-232 Standard. Основы передачи данных и интерфейс RS-232 [Электронный ресурс]. – Режим доступа: https://www.camiresearch.com/Data_Com_Basics/RS232_standard.html (дата обращения: 01.05.2025).
7. Postel J. User Datagram Protocol (RFC 768) [Электронный ресурс]. – Интернет-инженерный совет (IETF). – 1980. – Режим доступа: https://datatracker.ietf.org/doc/html/rfc768 (дата обращения: 01.05.2025).

УСТРОЙСТВО ЦИФРОВОГО МОНИТОРИНГА ТЕМПЕРАТУРЫ

ПАНФИЛОВА Е.Д.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Датчики температуры с ШИМ-выходом позволяют осуществить мониторинг температуры в устройствах. В работе рассмотрена синтезируемая схема обработки сигнала датчика и построена на практике в среде САПР QUARTUS II. Рассчитаны частоты работы. Схема была протестирована, результаты работы соответствуют ожидаемым значениям.

Ключевые слова: датчики температуры, датчики с ШИМ выходом, САПР QUARTUS II, частота генератора.

Введение

Мониторинг температуры в системах автоматизированного контроля работоспособности оборудования важен, так как позволяет обеспечивать надежность его работы. Использование датчика температуры позволит снизить риск отказов оборудования и продлить срок службы приборов.

Измерение температуры может производиться датчиками разных типов – терморезисторами, термопарами, датчиками на основе p-n переходов [1]. В цифровых системах наиболее удобно выполнять мониторинг с использованием датчиков с квазицифровым широтно-импульсно модулированным (ШИМ) выходом.

Выбор датчика температуры

Для работы был выбран датчик с ШИМ-выходом, поскольку он, с одной стороны, формирует непрерывный поток импульсных входных сигналов, что позволяет осуществлять мониторинг температуры, и при этом сигнал может быть обработан на цифровых элементах.

Работа датчика с ШИМ-выходом начинается с преобразования температуры в напряжение. Внутренний сенсор создает напряжение, пропорциональное абсолютной температуре. Далее полученное напряжение сравнивается с опорным напряжением в компараторе.

Результат, полученный после сравнения (1 или 0), поступает на схему дельта-сигма(Δ δ) [2], которая состоит из интегратора, компаратора и однобитного цифро-аналогового преобразователя. Основная часть схемы – интегратор, который накапливает значение ошибки, то есть разницы между текущем значением и выходом из однобитного ЦАП (цифро-аналогового преобразователя). Если ошибка положительна, значение интегратора возрастает, иначе уменьшается. На выход интегратора аналоговое напряжение.

Выход интегратора подается на компаратор, где значение сравнивается с 0. Если выход положительный, компаратор выдает 1, иначе 0. Далее выход компаратора подается на 1-битный ЦАП, который выдает опорное напряжение, если выход компаратора 1, иначе 0. Так, цифровой сигнал преобразуется обратно в аналоговый для подачи в интегратор.

Датчики, работающие по такому принципы производятся компаниями Analog Device [3, 4], Maxim [5], Melexis [6].

В представленной работе использовался датчик TMP03/TMP04 [3], в котором выходной сигнал представляется в виде ШИМ сигнала с постоянной длительностью единичной

составляющей импульса и переменной скважностью, а итоговая температура высчитывается по формуле

$$t = 235 - \frac{400 \cdot T_1}{T_2}, \quad (1)$$

где T_1 –длительность единичного значения импульса, T_2 – длительность паузы. Вид выходного сигнала представлен на рис. 1. Максимальная длительность $T_1 = 12$ мс, а $T_2 = 44$ мс (при максимальной температуре 125℃). Погрешность такого датчика 1,5 ℃.

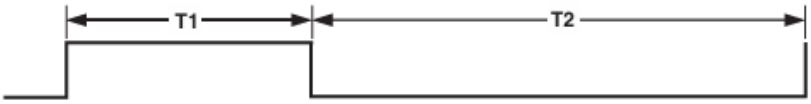


Рис. 1. Выходной сигнал датчика

Существуют датчики с более высокой точностью, например, TMP05/TMP06 [4], их погрешность составляет 0,5 ℃. При этом они отличатся формулой для расчета температуры:

$$t = 419 - \frac{768 \cdot T_1}{T_2},$$

сохраняя все особенности работы, присущие рассмотренному типу датчиков.

Проектирование контроллера датчика температуры

Для обработки сигнала датчика TMP03/TMP04 была разработана схема, основанная на подсчете длительности выходного сигнала с последующим арифметическим расчетом. Схема была реализована в САПР Quartus II [7] и выполнен ее синтез. RTL-вид схемы представлен на рис.2.

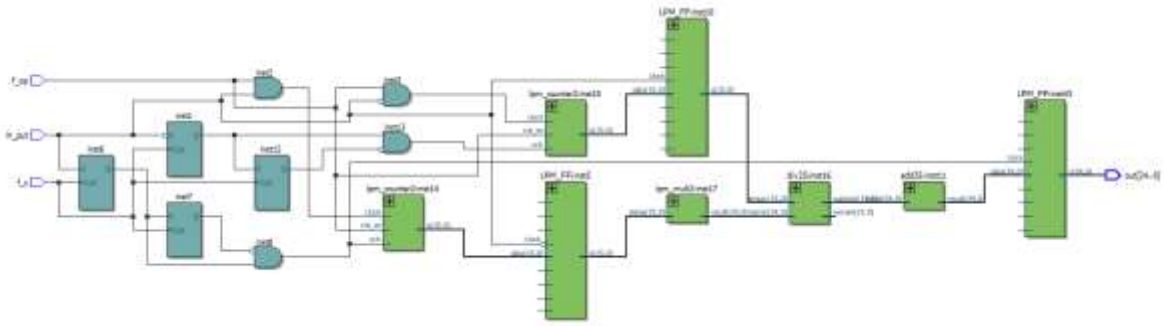


Рис. 2. RTL-вид синтезированной схемы

При расчете опорной частоты работы схемы учитывалась, что максимальная погрешность количества импульсов при подсчете значений T_1 и T_2 составляет ±2 импульса. Расчет частоты произведен следующим образом:

$T = \frac{N}{f}$ или $\frac{T_1}{T_2} = \frac{N_1}{N_2}$, где f – частота, N, N_1, N_2 – количество импульсов, соответствующих периоду T, а также значениям T_1 и T_2.

Погрешность датчика составляет 1,5 ℃, то есть

$$\Delta\frac{N_1}{N_2} \leq 0{,}00375$$

Таким образом, необходимо выбрать частоту, чтобы ошибка ±2 импульс не давала погрешность отношения, превышающую 0,00375. Чем больше количество импульсов, тем

меньше влияние ошибки. Стоит учитывать, что N_1 почти в 3,5 раз меньше, чем N_2, значит влияние ошибки на N_1 будет больше, чем на N_2.

При частоте 20 кГц (T_1 = 12 мс, T_2 = 43 мс):

$$N_1 = T_1 \cdot f = 12 \cdot 10^{-3} \cdot 20 \cdot 10^3 = 240 \text{ (импульсов)}$$

$$N_2 = T_2 \cdot f = 43 \cdot 10^{-3} \cdot 20 \cdot 10^3 = 860 \text{ (импульсов)}$$

$$\frac{N_1}{N_2} \approx 0{,}279$$

$$t = 235 - 400 \cdot \frac{N_1}{N_2} \approx \ 123{,}4$$

Худший случай: при ошибках -2 импульс N_1 и +2 импульс N_2. Так, N1 = 238, N_2 = 862.

$$t_н = 235 - 400 \cdot \frac{N_1}{N_2} \approx \ 123{,}6$$

$$\Delta t = |t - t_н| = 1{,}15 \text{ ℃}$$

Даже в худшем случае ошибка не превысила 1,5 ℃ . Значит, частота, достаточная для обеспечения корректной работы, должна быть не менее 20 кГц.

Для подсчета длительности единичного и нулевого значений ШИМ сигнала используются счетчики: в первом фиксируются значения, соответствующие единичному значению, а во втором – нулевому. Процесс записи управляется сигналом от датчика с использованием импульсов опорной частоты, коммутируемых через логическое «И». Результаты подсчета в конце каждого временного отрезка фиксируются в 12-ти разрядных регистрах.

Так, для корректной работы счетчиков можно определить диапазон опорной частоты: 20 кГц – 70 кГц.

Выходной сигнал датчика также будет использован как сигнал управления записи в оба регистра для того, чтобы последнее на момент окончания соответствующего временного отрезка значение каждого счетчика было сохранено.

Первый регистр количество импульсов, соответствующих единичному значению выходного сигнала (T_1), а второй – нулевому значению (T_2). Согласно формуле (1), данные из первого регистра умножаются на константу 400 с помощью элемента умножения. Результат умножения делится на содержимое второго регистра (T_2) с использованием делителя. Результат деления вычитается из константы 235.

Для управления работой последовательных устройств предусмотрена повышенная частота. На основе этой частоты формируются короткие импульсы для управления счетчиками. Для этого использована классическая схема детектора фронта на основе сдвигового регистра. Выход первого триггера и инвертированный выход второго собираются логическим «И», и формируется короткий импульс. В схеме использованы два таких детектора: первый короткий импульс сбрасывает счетчик, который определяет длительность T_1, и фиксирует результат в регистре; второй короткий импульс формируется на основе инвертированного сигнала датчика, такой импульс сбрасывает счетчик, определяющий период T_2.

Анализ синтезируемой схемы показал, что арифметические операции были выполнены без использования встроенных арифметических вычислителей. Задействовано 718 логических ячеек, часть из них переведена в арифметический режим. Характеристики синтезируемой схемы приведены в таблице 1.

Частота сигнала тактирования триггеров должен быть больше частоты счетчиков. Для этого можно использовать дополнительный тактовый сигнал или повысить опорную частоту с использованием встроенного модуля PLL.

Таблица 1

Характеристики синтезируемой схемы

Компоненты	Количество
Ячейки	755
Ячейки комбинационных функций	718
Логические ячейки в нормальном режиме	352
Логические ячейки в арифметическом режиме	366
Регистры	93

Моделирование работы контроллера датчика температуры

Выполнено моделирование работы схемы. Результаты представлены на временных диаграммах на рис. 3. После первого периода выводится значение температуры, оно будет держаться до тех пор, пока не закончится следующий период и не выведется новое значение температуры.

Рис. 3. Временные диаграммы

Для тестирования были заданы T_1 = 14,5 мкс, T_2 = 30,2 мкс. Тестовый выход должен составлять 43,9℃, но составил 43℃. Во втором периоде T_1 = 14,7 мкс, T2 = 30,2 мкс, t = 40,3. Все тесты соответствуют точности датчика TMP03/TMP04.

Выводы

В ходе работы создан контроллер датчика температуры с ШИМ-выходом. Использован классический подход, основанный на заполнении сигнала датчика опорной частотой с последующими вычислениями. Получено экономичное решение, так как в схеме не задействованы встроенные арифметические вычислители.

Список литературы

1. Н. М. Сафьянников, О. И. Буренева, А. Н. Алипов. Информационно-измерительные преобразователи киберфизических систем : учебное пособие для вузов - 2-е изд., стер. - Санкт-Петербург : Лань, 2024. - 236 с.
2. Дельта-сигма модуляция: назад в будущее, А. Махлин. Компоненты и технологии. 2010. №11, стр. 154-158
3. Reference manual. TMP03/TMP04. Analog Devices
4. Reference manual. TMP05/TMP06. Analog Devices
5. Цифровые датчики температуры с ШИМ-выходом. [Электронный ресурс]. URL: https://www.sensorica.ru/d2-3d.shtml
6. Датчики температуры и влажности на микросхемах. [Электронный ресурс]. URL: https://rs-catalog.ru/temperature-and-humidity-sensor-ics/?features_hash=448-192229

7. Altera Corporation. Introduction to Quartus II.

ПРИМЕНЕНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ГЕМАТОЛОГИИ

РОГОЗИНСКАЯ А.Л.

Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО».

Аннотация. Рассмотрены уже существующие разработки аппаратных и программных средств для гематологии. Выделены основные направления дальнейшего развития и указаны основные препятствия, мешающие дальнейшему внедрению информационных технологий в гематологию.

Ключевые слова: гематология, машинное обучение, информационные технологии, медицина

Ведение

Актуальность темы состоит в потребности предупреждения вероятности возникновения человеческой ошибки и снижения рабочей нагрузки на медицинских работников.

Цель исследования состоит в рассмотрении перспектив применения информационных технологий в гематологии.

Для достижения поставленной цели необходимо решить следующую задачу: рассмотреть и сравнить уже существующие технологии.

Объектом исследования в данной работе являются информационные технологии, а методом исследования рассмотрение уже существующих технологий.

На сегодняшний день разработано множество способов внедрения аппаратных и программных инновационных технологий в работу медицинского персонала. Нововведения позволяют избежать человеческой ошибки, а также снизить потребность в обучении и найме большого количества работников. Хотя, о полной автономности в решении медицинских задач говорить рано.

Анализ крови необходим для постановки различных диагнозов от диабета до оникомикоза. Он доступен во многих лабораториях, стационарных и амбулаторных медицинских учреждениях. Существует несколько основных анализов крови: биохимический, гормональный, полный и на наличие инфекций. Результат представляет собой перечисление морфологии форменных элементов крови и указание их количества.

Использование машинного обучения

Рекуррентные нейронные сети в медицине наиболее часто используются для анализа данных электронных медицинских карт. Для обработки анализов можно использовать сверточные нейронные сети, применяемые в компьютерном зрении. Каким образом будет вестись обработка данных зависит от применяемой математической модели. Для успешной работы модели необходимы правильно подготовленные: обучающая выборка, валидационная выборка и тестовая выборка. Лучше использовать комбинированные данные, а не однородные.

В НМИЦ ДГОИ им. Дмитрия Рогачева успешно использовали программу резидента фонда «Сколково», компании «Платформа Третье Мнение». Точность исследований с

применением искусственного интеллекта составила более 95% [1]. Технология «Метод для выделения и классификации типов клеток крови с помощью глубоких свёрточных нейронных сетей» была запатентована ещё в 2020 году. Она умеет выделять клетки, характерные при злокачественных заболеваниях крови. Рост заболеваемости раком растёт, но смертность от онкозаболеваний снижается. Однако, рост заболеваемости раком показатель благоприятный, так как указывает на то, что заболевание чаще диагностируют. Зачастую результаты, полученные с помощью искусственного интеллекта должен утвердить врач. Компания «Платформа Третье Мнение» так же применяет искусственный интеллект для мониторинга пациентов, лучевой диагностики, стоматологии и офтальмологии.

Так же создаются экспертные системы для выявления возможных заболеваний крови. В статье на популярном ресурсе «Habr» врач-гематолог с рабочим стажем более четырнадцати лет рассказывает, как разработал подобное приложение на языке программирования R. R предназначен для статистической обработки данных и работы с графикой. Данная программа проста в реализации, заданы всего двадцать входных параметров, используются 4 алгоритма XGBoost. Данные для обучения (результаты общего анализа крови более двадцати двух тысяч пациентов) подобраны медицинским специалистом, а не молодым студентом IT-специальности, что повышает эффективность работы приложения [2].

В «Санкт-Петербургском национальном исследовательском университете информационных технологий, механики и оптики» в лаборатории «Компьютерных технологий» вместе с зарубежными коллегами из биотехнологической компании «Insilico Medicine» была создана система Aging.AI, способная определить по результатам базового анализа крови возраст человека. Применяется технология глубокого обучения. Модель построена на комбинированных данных, взятых из разных континентов. Подобная система так же, как и предыдущая находится в открытом доступе. Выделены ключевые биомаркеры, наиболее влияющие на старение. Биологический возраст человека может отличаться от хронологического. Подобная разработка нужна в медицине, например, для выявления риска развития деменции и в работе систем страхования [3].

В ведущем исследовательском Квинследском университете Австралии был создан микроскоп, помогающий определять заболевания крови с помощью искусственного интеллекта. Технология разрабатывалась много лет [4].

Анализатор крови

Помимо программных новшеств на помощь медицинским работникам приходят аппаратные нововведения. У большинства есть один существенный недостаток – высокая цена. Существуют открытые и закрытые биохимические анализаторы. Закрытые анализаторы используют ограниченное число реагентов, а открытые анализаторы используют реагенты многих производителей. В «Санкт-Петербургском национальном исследовательском университете информационных технологий, механики и оптики» 2015 году А. А. Маргун, Д. Н. Базылев, К. А. Зименко, А. С. Кремлев разработали структуру и элементную базу анализатора крови закрытого типа [5]. Были минимизированы габаритные размеры прибора. Хемилюминесцентный метод используется для работы с результатами клинического анализа. Шаговые двигатели используются как исполнительные устройства.

Заключение

В заключении необходимо сказать, что дальнейшие перспективы применения информационных технологий в гематологии зависят от финансирования отрасли. Двигателем развития является нужда в высокой скорости обработки информации. Необходимо, что бы разработками занимались совместно медицинские работники и инженеры, дополняя знания и умения друг друга. Стоит отметить, что наиболее популярной и полезной технологией в России является «метод для выделения и классификации типов клеток крови с помощью глубоких свёрточных нейронных сетей» компании «Платформа Третье Мнение».

Литература

1. Медвестник. Цифровая медицина. Точность обработки мазка крови искусственным интеллектом достигла 97%. [Электронный ресурс]. URL: https://medvestnik.ru/content/news/Tochnost-obrabotki-mazka-krovi-iskusstvennym-intellektom-dostigla-97.html (Дата обращения: 16.11.2024).
2. Хабр. Как врач сделал себе ИИ помощника. Часть 2. [Электронный ресурс]. URL: https://habr.com/ru/articles/710154 (Дата обращения: 16.11.2024).
3. Хабр. Нейросеть, определяющая возраст по анализу крови, — разработки ученых Университета ИТМО. [Электронный ресурс]. URL: https://habr.com/ru/companies/spbifmo/articles/350624/ (Дата обращения: 17.11.2024).
4. RG RU. Австралийские исследователи впервые применили искусственный интеллект для анализа крови в лабораториях. [Электронный ресурс]. URL: https://rg.ru/2023/07/21/avstralijskie-issledovateli-vpervye-primenili-iskusstvennyj-intellekt-dlia-analiza-krovi-v-laboratoriiah.html (Дата обращения: 01.12.2024).
5. Маргун А. А., Базылев Д. Н., Зименко К. А., Кремлев А. С. Структура и элементная база анализатора крови закрытого типа // Изв. вузов. Приборостроение. 2015. Т. 58, № 10. С. 859—861.

РАЗРАБОТКА АЛГОРИТМА ЛОКАЛИЗАЦИИ ИСТОЧНИКА ЗАГРЯЗНЕНИЯ РАСПРЕДЕЛЕННЫМ РОБОТОТЕХНИЧЕСКИМ КОМПЛЕКСОМ НА ОСНОВЕ МУЛЬТИАГЕНТНОГО ПОДХОДА

ХАСАНОВ А.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В статье рассматривается реализация алгоритма для решения задачи локализации одиночного источника загрязнения воздуха на заданной территории. В работе алгоритма применяется мультиагентный подход, который позволяет выполнять задачу с помощью группы элементов в составе единого робототехнического комплекса в автономном режиме. Результаты разработки алгоритма были проверены путем проведения испытаний посредством имитационного моделирования в специально разработанном программном пакете.

Ключевые слова: Мультиагентная система, роботехнические комплексы, автономные интеллектуальные агенты, загрязнение воздуха, локализация источников загрязнения

1. Анализ существующих методов локализации источника загрязнения

Решение задачи локализации с применением мобильных робототехнических комплексов (мобильных РТК) рассматривается исследователями с начала 1990-х годов [1]. Из-за ряда недостатков использование на сложной местности большой площади являлось невозможным. С развитием технологий производства – использование таких систем стало экономически целесообразным [2]. На данный момент наиболее подходящим робототехническим комплексом для выполнения таких задач являются беспилотные комплексы. Авторы различных исследований уже использовали беспилотные комплексы для решения задачи поиска загрязнений окружающей среды [3 – 5], однако они обладают рядом недостатков, связанным в первую очередь с низкой автономностью описываемых систем и сильной зависимостью от априорных данных. Одним из эффективных способов устранения данных недостатков является использование мультиагентного подхода [6, 7], позволяющего добиться высокой степень автономности, гибкости к условиям и универсальности системы для различных конфигураций территории и загрязнения [8].

2. Методы, используемые для реализации алгоритма локализации

1. Построение модели окружающей среды

В данной модели территория представляет собой двумерную ортогональную координатную плоскость в виде квадрата, состоящую из точек. Имея в виду, что территория представляет собой множество *P,* состоящее из элементарных участков $M = \{p\}$ – каждый из них характеризуется некоторым набором параметров (координаты, значение концентрации, направление ветра и т.д.). Это позволяет рассматривать её как объект, взаимодействовующий с другими объектами, в частности, с моделью загрязнения и группой агентов. Тогда набор объектов точек с ненулевым параметром концентрации загрязняющего вещества – это уже множество «загрязнения», а набор объектов точек, на которых в данный момент времени находятся агенты – это уже множество «группы агентов».

Для описания загрязнения в данной работе используется модифицированная модель турбулентной диффузии, определяющая распространение газовоздушной примеси в атмосфере. Непосредственно характеристики модели и её модификация были описаны в [9]. Она является типовой и входит в перечни допустимых, сформированных уполномоченными ведомствами РФ и США. В ней рассматривается так называемый приземный слой атмосферы (соответствует высоте 2 метра над землей). Согласно модели, загрязнение представляет собой множество точек с относительными величинами концентраций некоторого загрязняющего вещества. Распределение характеризуется двумя главными компонентами – источником загрязнения и его следом (факелом). Для факела значение концентрации является функцией от координат и параметров источника загрязнения.

$$c_p(x, y) = s_2(y) \cdot \forall x_i \in \{x_0; x_n\}(s_1(x))$$

где $s_1(x)$ и $s_2(y)$ – это безразмерные функции, определяющие форму следа загрязнения в зависимости от его параметров [13].

Для целей работы конкретные параметры источника не имеют влияния на работу алгоритма. На рисунке 1 изображена итоговая модель.

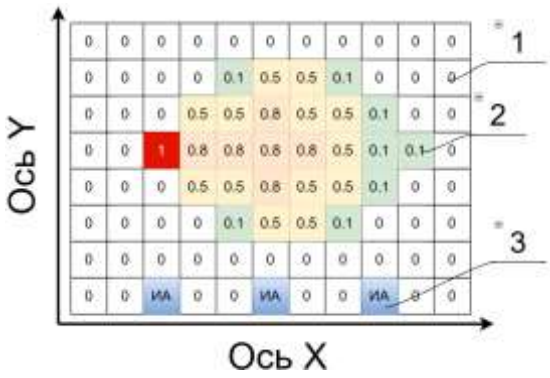


Рис 1. Модель системы, где 1 – территория, 2 – загрязнение, а 3 – группа интеллектуальных агентов

Было предусмотрено, что детализация, форма и размер модели загрязнения может регулироваться специальными гиперпараметрами, описание которых будет опущено.

2. Применение мультиагентного подхода в построении модели работы группы

Разработанная модель окружающей среды представляется набором точек пространства. Перемещаясь по точкам пространства – агенты исследуют его в поиске той точки, в которой относительная концентрация соответствует источнику загрязнения.

Примем допущение, что в каждый момент времени в точке пространства может находиться только какой-то один агент. Тогда подмножество точек территории, на которых находятся агенты назовем группой агентов.

Применение мультиагентного подхода предполагает, что каждый из интеллектуальных агентов должен обладать автономностью во время выполнения общей миссии. Для этого каждый из них должен обладать: набором уникальных параметров, средствами коммуникации, средствами измерения для исследования окружающей среды, органами управления, а также политикой функционирования, определяющей цели миссии группы. Тогда задача описания группы сводится к описанию индивидуального агента.

Каждый из агентов обладает собственным набором параметров в виде вектора. А процедура обмена векторами параметров между агентами назовем «коммуникацией» внутри группы. Помимо коммуникации друг с другом, в рамках алгоритма агенты должны получать информацию о состоянии окружающей среды. Поэтому во время исследования территории, каждый из них используя датчики собирает следующую информацию (проводит измерения):

• $c_p(x, y)$– Значение относительной величины концентрации в текущей точке;

• $d_p(x, y)$ – Значение направления скорости ветра в текущей точке;

Тогда процесс принятия решения каждым из агентов основан на получении информации извне и её последующей обработке. Для корректной работы необходимо чтобы каждый из агентов выполнял итерации синхронно. Это позволяет описать миссию мультиагентной системы в виде целевой функции T следующего вида:

$$T\{A\} = f\left(p_5 \rightarrow true;\ t_n \rightarrow min\right)$$

где t_n – количество временных итераций, требуемых для смены флага нахождения источника на *true*, p_5 –один из параметров в векторе, флаг окончания миссии.

В данной итерации исследования алгоритм основан на передаче между агентами информации о «режимах» функционирования. Поскольку задача не сводится к слепому поиску конкретной точки из множества, а основана на кооперативном исследовании территории – как только какой-то из агентов в группе попадает хотя бы на след загрязнения

– режим работы всей группы целиком меняется и тем самым использует информацию о найденной «зацепке» для более быстрого определения источника.

В начале миссии все агенты находятся в *режиме обхода* и перемещаются по территории с помощью карты весовых коэффициентов. Весовые коэффициенты для каждой из точек пространства рассчитываются в зависимости от расположения агентов на территории и расстояний от каждого из них до рассматриваемой точки пространства.

В зависимости от дальнейшего хода исследования территории, агенты могут сменить своё состояние на два других режима:

• *Режим локализации.* Активируется для агента в момент получения информации от датчиков о том, что в точке была обнаружена минимальная пороговая концентрация загрязняющего вещества $c_p(x_c, y_c)$. При его активации агент ориентируется на данные от анемометра (направление ветра) и двигается в противоположную ему сторону. Если в следующей точке концентрация не меньше – он продолжает следование по этому направлению до тех пор, пока не обнаружит точку с источником.

• *Режим следования*. Активируется в момент получения агентом информации о том, что другой агент сменил своё состояние на «режим локализации». После чего агент начинает движение в его сторону. При этом, если во время следования агент попадает на точку с ненулевым загрязнением – он тоже переходит в режим «локализации».

Тогда общий алгоритм с применением мультиагентной системы следующий:

1. Расчёт взвешенного графа $G(a_i)$ для заданной начальной конфигурации территории и группы интеллектуальных агентов;
2. Обход территории вглубь для оптимального исследования территории за наименьшее время;
3. Локализация источника загрязнения всей группой агентов совместно, путем передачи друг-другу информации о собственном состоянии и данных измерений;

3. Проведение имитационного моделирования и его результаты

Для проверки работоспособности алгоритма был разработан программный пакет на языке Python с использованием библиотек Matplotlib и NumPy. Так же он позволяет конфигурировать каждую часть модели и отображать результаты выполнения алгоритма в текстовом и графическом виде.

Алгоритм выполняется итерациями в виде цикла действий: сбор информации, принятие решения о следующей точке, перемещение, проведение измерений, смена состояния, передача информации другим агентам. Это позволяет обеспечить синхронность и параллельность выполнения миссии каждым из агентов и превратить метрику, определяющую время выполнения миссии в полностью безразмерную для более объективной оценки алгоритма.

Мерой качественной оценки работы алгоритма является время выполнения t, равное количеству итераций от начала до конца миссии. Для исследования изменения скорости выполнения алгоритма были проведены опыты с использованием различных конфигураций основных параметров модели. Для каждого из них загрязнение было идентичным и находилось в одном месте. Результаты представлены в таблице 1.

Таблица 1 – Результаты экспериментов при различных конфигурациях

Число агентов, *n*	**Скорость выполнения алгоритма, итерации *t***				
	***S* = 50 (малая)**	***S* = 75 (средне-малая)**	***S*= 100 (средняя)**	***S*= 150 (средне-большая)**	***S* = 200 (большая)**
1	495	1167	2266	5191	9223
2	344	938	1820	4652	8591

3	237	556	1031	2524	4549
4	**180**	**382**	**635**	**1846**	**3334**
5	179	401	878	1923	4324

Видно, что для эксперимента, проведенного в описанных условиях – прирост в скорости приостанавливается на размере группы в 4 агента. Такой результат можно объяснить тем, что при текущей версии алгоритма затраты на обеспечение коммуникации и поддержания работоспособности системы приводят к большей величине издержек, чем оптимизации. Аналогичная проблема наблюдается, например, в проблеме многопоточности, когда достигается зона, в которой обслуживание системы занимает больше ресурсов, чем позволяет сэкономить.

Такой результат также согласовывается с решением различных задач по инженерной реализации заданной системы. Такие задачи как обеспечение коммуникации, синхронности и взаимодействия агентов друг с другом тем сложнее, чем большее количество в группе следует связать друг с другом. Поскольку каждый из агентов является полностью самостоятельным и автономным – такая система требует больших трудозатрат для обеспечения её работоспособности.

Список литературы

1. Rozas R., Morales J., Vega D. Artificial smell detection for robotic navigation // Fifth International Conference on Advanced Robotics Robots in Unstructured Environments. – IEEE, 1991. – P. 1730-1733.
2. Marques L., Martins A., de Almeida A. T. Environmental monitoring with mobile robots // 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems. – IEEE, 2005. – P. 3624-3629.
3. Bolla G. M. ARIA: Air pollutants monitoring using UAVs / G. Bolla, M. Casagrande, A. Comazzetto [et al.] // 2018 5th IEEE International Workshop on Metrology (MetroAeroSpace). – IEEE, 2018. – P. 225-229.
4. Francis A. Gas source localization and mapping with mobile robots: A review / A. Francis, S. Li, C. Griffiths [et al.] // Journal of Field Robotics. – 2022. – Vol. 39. – N 8. – P. 1341-1373.
5. Zhou B. A review of Space-Air-Ground integrated remote sensing techniques for atmospheric monitoring / B. Zhou, S. Zhang, R. Xue [et al.] // Journal of Environmental Sciences. – 2023. – Vol. 123. – P. 3-14.
6. Wang Y. Pollution source localization in complex environments using a swarm of mobile agents/Y.Wang, Y. Hu, J Wang [et al.] // IEEE Sensors Letters. – 2021. – Vol. 6. – N 1. – P. 1-4.
7. Bolla G. M. ARIA: Air pollutants monitoring using UAVs / G. Bolla, M. Casagrande, A. Comazzetto [et al.] // 2018 5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace). – IEEE, 2018. – P. 225-229.
8. Викснин И. И., Лазарев Э. А., Хасанов А. Локализация источника загрязнения воздуха на заданной территории с применением мультиагентного подхода для управления группировкой беспилотных летательных аппаратов // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2023). – СПб.: Издательство СПбГУТ, 2023. – С. 470-475.
9. Khassanov A. M., Suloeva E. S. The Modeling of Carbon Emissions from a Point Source // 2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). – IEEE, 2022. – P. 1534-1537.

РОБОТ-СБОРЩИК КУБИКА РУБИКА

В.С. Хорошков, Т.Е. Казакова, А.В. Колесникова

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В статье описывается процесс создания робота для автоматизированной сборки кубика Рубика: проектирование 3D-модели устройства, выбор аппаратных компонентов, трассировка

и изготовление печатной платы, реализация алгоритмов распознавания цвета и решения кубика Рубика. Приводится анализ неудачных конструкторских решений, возникших в процессе разработки.

Ключевые слова: кубик Рубика, робототехника, машинное зрение, машинное обучение.

Введение

Кубик Рубика – это механическая головоломка в форме куба. Каждую грань куба, разделенную на 9 квадратов одного из шести цветов, можно вращать. Цель головоломки – путем поворота граней привести куб в такое состояние, чтобы каждая из граней содержала квадраты только одного цвета. Эта головоломка сочетает в себе кажущуюся простоту и колоссальную вычислительную сложность, ведь при двенадцати возможных действиях – поворотах одной из граней, может получиться более 43 квинтиллионов комбинаций.

Эволюция методов решения кубика Рубика развивалась от интуитивных поисков первых энтузиастов до современных высокоэффективных алгоритмов, способных находить кратчайшую последовательность действий. В момент появления головоломки в 1974 году на ее решение могли уходить месяцы. На сегодняшний день обычный человек, умеющий собирать кубик, тратит на решение примерно от 1 до 8 минут. Спортсмены-спидкуберы собирают кубик гораздо быстрее – нынешний рекорд в 3.05 секунды принадлежит семилетнему Сюаньюю Гэн [1].

Доказано, что для классического кубика Рубика, из любой его конфигурации, существует решение до 20 ходов, которое приводит кубик в собранное состояние [2]. Простые человекоориентированные методы, такие как послойная сборка [3] и CFOP [4], делят процесс на этапы, понятные человеку, и не требуют сложных вычислений. Такие алгоритмы позволяют решать головоломку за 60–250 ходов. Метод послойной сборки является базовым 7-этапным алгоритмом для начинающих. В отличие от него метод CFOP требует запоминания 119 комбинаций, при этом позволяет сократить количество ходов и добиться большей скорости сборки.

Продвинутые алгоритмы используют нетривиальный анализ состояний (ходов) и потому пригодны только для использования на ЭВМ. К таким алгоритмам можно отнести двухфазный алгоритм Коцембы [5]. Алгоритм разбивает задачу на 2 фазы, в каждой фазе ограничиваются доступные ходы, тем самым сокращая группу возможных состояний кубика Рубика. Это снижает количество комбинаций в каждой группе с 43 квинтллионов до ~20 млн состояний в фазе 1 и ~10 тысяч в фазе 2. Особый интерес представляют гибридные методы машинного обучения, такие как DeepCubeA [6], сочетающий глубокие нейронные сети с алгоритмами поиска пути.

В конце 2010-х появились первые роботы, собирающие головоломку автоматизированным способом. Энтузиасты конструировали робота из набора Lego mindstorms EV3, который мог считывать цвета, вращать одну грань и поворачивать кубик [7]. Такой подход занимал несколько минут. Позднее появились роботы, способные вращать все шесть граней [8]. Роботы-сборщики могут по-разному определять начальное положение кубика: некоторые требуют ручного ввода параметров [8]; другие имеют камеры с нескольких сторон [9] и определяют цвет с помощью машинного зрения; могут использоваться умные кубики, оборудованные встроенными датчиками и передающие текущее состояние по Bluetooth прямо на управляющее устройство [10].

Развитие роботов-сборщиков кубика Рубика отражает общий технический прогресс: от медленных первых прототипов, собирающих головоломку за 5–10 минут, до современных сверхбыстрых систем. Текущий рекорд по сборке составляет

0.203 секунды [9]. Предыдущее достижение в 0.305 секунды [11] было улучшено благодаря оптимизации поворота – в старом рекорде несколько миллисекунд тратилось на стабилизацию грани.

Далее в статье будет последовательно описан процесс создания подобного робота: моделирование конструкции, её сборка, прототипирование платы, алгоритмическая составляющая и её реализация.

Конструкция робота

Основу робота составили шесть шаговых двигателей NEMA 17HS4401, управляемых драйверами A4988 с помощью микроконтроллера Arduino Due. Для распознавания цветов граней использовалась веб-камера Logitech C920. После нахождения решения ЭВМ передает команды микроконтроллеру, генерирующему управляющие импульсы для шаговых двигателей.

Для робота-сборщика кубика Рубика была разработана конструкция на основе алюминиевых профилей, представленная на Рис. 1, обеспечивающих необходимую жёсткость. Чтобы упростить доступ к кубику, боковые двигатели закрепляются на быстросъёмных основаниях – это позволяет легко извлекать кубик без потери точности позиционирования, сохраняя при этом модульность конструкции для обслуживания. Жёсткое крепление не требуется, так как небольшой люфт компенсируется автоматически при последующем повороте грани.

Одной из ключевых задач стало создание механизма передачи вращения с вала двигателя на грань кубика. На Рис показаны этапы эволюции переходника, демонстрирующие последовательную оптимизацию решения. Версия 1 выполнена с прямыми углами и оказалась неподходящей из-за механического заклинивания соседних граней при вращении. В версии 2 наличие скругленных краев и усиленных зацепных элементов, привело к недопустимому увеличению рабочего зазора между элементами кубика. Версия 3 не обеспечивала требуемой механической надёжности – уменьшенные зацепные элементы допускали холостое прокручивание без передачи вращения на грань. Финальная версия 4 представляла собой переходник, полностью заменяющий съёмную крышку центрального элемента (у современных кубиков Рубика под ней находится настроечные винты кубика, возможность съема предусмотрена изначально). Верхняя часть переходника версии 4 в точности повторяет геометрию оригинальной крышки.

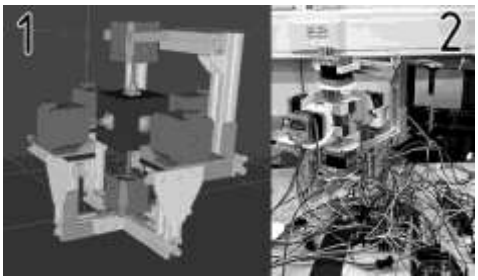


Рис. 1. Конструкция робота-сборщика. 1 – 3D-модель. 2 – Физическая реализация

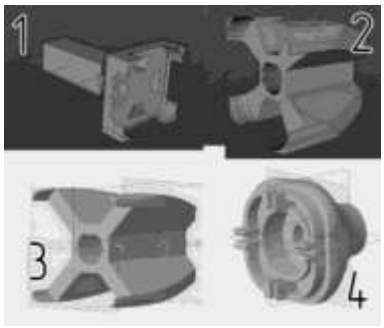


Рис. 2. Эволюция переходников с вала двигателя на грань кубика. 1 – Первый вариант: захват центральной клетки. 2 – Второй вариант: добавлены скругленные края захвата. 3 – Третий вариант: захват еще меньше и плавнее. 4 – Итоговый вариант: захват не зацепляет клетку, а заменяет ее.

Первоначальная сборка схемы на макетной плате, изображенная на Рис. 1, позволила протестировать работоспособность схемы**Ошибка! Источник ссылки не найден.**, но из-з а большого количества проводов конструкция выглядела громоздкой, а соединения оставались ненадёжными, кроме того, такая компоновка затрудняла обслуживание. Для устранения этих недостатков была разработана печатная плата, изображенная на **Ошибка! И сточник ссылки не найден.**, в среде проектирования EasyEDA. Далее выполнено ЛУТ-изготовление платы с механическим сверлением отверстий и ручной пайкой компонентов.

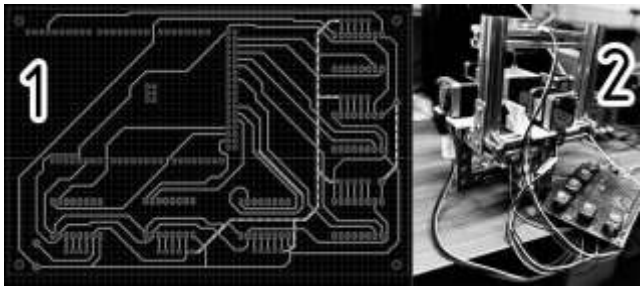


Рис. 3. Спроектированная схема печатной платы и её внедрение в модель робота. 1 – Итоговая схема платы, пунктиром изображен нижний слой. 2 – Общий вид сборщика со всеми компонентами.

Алгоритм определения состояния

Задача классификации цветов клеток кубика Рубика сводится к анализу усреднённого цвета каждого сегмента на изображении с использованием матричных преобразований. Сначала определяются границы кубика, к изображению каждой грани применяется аффинное преобразование для получения неискаженной перспективы, затем грань делится на 9 равных сегментов, для каждого из которых вычисляется средний цвет в RGB-пространстве. В данном проекте была использована тривиальная классификация на основе порогов для отношений цветовых каналов.

Алгоритм решения

Для решения кубика Рубика было использовано три различных алгоритма: послойной сборки, алгоритм Коцембы и нейросетевой алгоритм DeepCubeA [6]. Первые два алгоритма были реализованы на языке Java, нейросетевое решение получалось путем обращения к api сайта [12]. В таблице 1 для каждого алгоритма указано среднее время решения произвольной конфигурации, среднее количество ходов, приводящих к решению, и среднее время сборки кубика, представляющее собой суммы времени на решение и времени

на повороты. Послойная сборка оказалась самой быстрой в поиске решения, но наименее эффективной по количеству ходов. DeepCubeA показал максимальную оптимизацию шагов, однако потребовал значительно больше времени. Алгоритм Коцембы продемонстрировал оптимальное соотношение временных затрат и операционной эффективности.

Таблица 1

Сравнение алгоритмов

Алгоритм	Среднее время решения	Среднее количество ходов	Оценка среднего времени сборки
Алгоритм послойной сборки	0,59 мс	206,1	95 837 мс
Алгоритм Коцембы	1055,80 мс	41,2	20 235 мс
Алгоритм DeepCubeA	28207,97 мс	30,3	42 297 мс

Программная часть

Программный код для управления робота был написан на языке программирования Java. Актуальная версия находится в репозитории: https://github.com/vad-ik/RobotCubSolver

Заключение

Спроектированный и собранный робот успешно выполняет задачу сборки ркубика Рубика и может быть использован в образовательных целях для визуализации нетривиальных алгоритмов сборки, а также имеет потенциал для последующего участия в соревнованиях автоматической сборки на время.

Список литературы

1. Рекорды // Всемирная Ассоциация Кубика. URL: https://www.worldcubeassociation.org/results/records?show=mixed (дата обращения: 25.04.2025).
2. Rokicki T., Kociemba H., Davidson M., Dethridge J. The diameter of the Rubik's cube group is twenty // SIAM Review. 2014. Vol. 56, № 4. P. 645-670.
3. Marshall P. The Rubik's Cube: A Puzzle Toy Phenomenon // Journal of Recreational Mathematics. 1981. Vol. 14, № 3. P. 194-201.
4. Fridrich J. My speed cubing page [Электронный ресурс]. URL: https://www.speedsolving.com/wiki/index.php/Jessica_Fridrich (дата обращения: 25.04.2025).
5. Kociemba H. Optimal Cube Solver [Электронный ресурс]. URL: https://kociemba.org/cube.htm (дата обращения: 27.04.2025).
6. Agostinelli F. et al. Solving the Rubik's cube with deep reinforcement learning and search // Nature Machine Intelligence. 2019. Vol.1, № 8. P. 356-363. DOI: 10.1038/s42256-019-0070-z.
7. Робот, собранный из набора Лего // MindCuber.com [Электронный ресурс]. URL: https://mindcuber.com/mindcub3r/mindcub3r.html (дата обращения: 25.04.2025).
8. Робот, который вращает все шесть граней, но медленнее человека // URL: https://digitalocean.ru/n/rubik-robot (дата обращения: 25.04.2025).
9. История робота, который быстрее всех в мире собирает кубик Рубика // Habr [Электронный ресурс]. URL: https://habr.com/ru/companies/yandex/articles/831708/ (дата обращения: 30.04.2025).
10. Описание самособирающегося кубика рубика // makezine.com [Электронный ресурс].
11. URL: https://makezine.com/projects/smart3-the-self-solving-rubiks-cube/ (дата обращения: 25.04.2025).
12. Информация о роботе Мицубиши // URL: https://hightech.fm/2024/05/27/mitsubishi-robot-rubik (дата обращения: 25.04.2025).
13. Solve the Rubik's Cube // URL: https://deepcube.igb.uci.edu/ (дата обращения: 25.04.2025).

ТЕХНОЛОГИЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ВИЗУАЛЬНОГО КОНТРОЛЯ КАЧЕСТВА ЭЛЕКТРОТЕХНИЧЕСКИХ ИЗДЕЛИЙ

П.В. Черникова

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В статье приведена технология автоматизированного визуального контроля качества электротехнических деталей малого размера для использования на крупносерийном производстве. Предложена модульная архитектура, реализующая анализ дефектов на основе сравнения изображений контролируемой детали с эталонами. Проведен эксперимент для определения факторов, влияющих на качество анализируемых снимков. Выявлены ограничения системы и определены возможности ее улучшения.

Ключевые слова: автоматизированный контроль качества, визуальный контроль, машинное зрение, поиск аномалий, электротехнические изделия

Введение

Развитие технологического прогресса приводит к росту объемов крупносерийного производства электротехнических изделий, при котором должны сохранятся требования к точности и стабильности контроля качества на каждом этапе. Классическим методом выполнения этой задачи является ручная проверка деталей оператором с помощью микроскопов и монокуляров. Это ограничивает производительность, иногда может приводить к ошибкам вследствие влияния человеческого фактора, а также требует затрат на обучение и оплату труда кадров [1-2]. Автоматизация процесса визуального контроля качества может решить обозначенные проблемы, однако её внедрение сопряжено с рядом трудностей, связанных со спецификой строения анализируемых деталей.

Условия автоматизации

При решении задачи автоматизации должны учитываться следующие особенности электротехнической продукции. Изделия имеют малые размеры, сложную и неустойчивую форму, а также могут состоять из различных материалов, таких как сталь, пластик, прозрачное или темное стекло. Эти факторы влекут трудности с фиксацией деталей, выбором условий освещения и съемки. Существующие промышленные системы автоматической визуальной оценки качества разнообразных малогабаритных деталей электроники используют широкий круг аппаратных средств: камеры машинного зрения, гиперспектральные камеры, рентген-установки и специализированные осветительные модули [3-5]. Конфигурации оборудования зависят от характеристик конкретных изделий, при этом такие решения зачастую основываются на использовании нейросетевых алгоритмов анализа, требующих сбора больших объемов входных данных, специфичных для конкретной категории деталей, а также значительных вычислительных и временных затрат на обучение и дообучение моделей. Это ограничивает универсальность и ценовую доступность применения готовых реализаций.

Экспериментальное исследование

С целью изучения влияния различных факторов на качество снимков, используемых алгоритмами обработки в процессе визуального выявления дефектов на изделиях, был проведен эксперимент, учитывающий всевозможные комбинации следующих параметров:

1. материал изделия:

   a. матовый, блестящий или окрашенный металл;
   b. прозрачное, мутное светлое или тёмное стекло;
   c. пластик;
2. характеристики освещения:
   a. направленный или рассеянный свет;
   b. расположение источников освещения относительно объектов;
   c. мощность изучения;
   d. совмещение с поляризационными антибликовыми фильтрами;
   e. освещение деталей «на просвет» или отражённое;
3. фон:
   a. цвет;
   b. материал;
   c. текстура;
4. оптическое увеличение.

В результате было установлено, что невозможно выделить универсальную для всех типов деталей комбинацию, которая бы обеспечивала одинаково стабильное качество анализа. Это подтверждает необходимость адаптации условий захвата изображений под особенности конкретных типов изделий и производственного процесса. Точные размеры деталей и расстояние от них до объектива камеры в промышленной установке будут определять такие параметры, как область обзора и рабочее расстояние устройства съемки. В свою очередь они будут влиять на выбор разрешения, увеличения и фокусного расстояния камеры.

Минимальное разрешение снимков камеры определяется согласно формуле на рисунке 1 [6].

$$R_{w,h} = \frac{S_{w,h}}{S_{\min}} \cdot N_{px} \qquad (1)$$

Рис. 1. Расчет минимального разрешения изображений по ширине или высоте

Где:

$S_{w,h}$ – размеры объекта по ширине или высоте (мм),

$S_{\min}$ – минимальный размер обнаруживаемой аномалии (мм),

N_{pix} – минимально необходимое количество пикселей для обнаружения одной аномалии.

Оптическое увеличение объектива определяется согласно формуле на рисунке 2 [7].

$$PMAG = \frac{S_{sensor}}{FOV} \qquad (2)$$

Рис. 2. Расчет оптического увеличения объектива

Где:

S_{sensor} – размер стороны сенсора камеры (мм),

FOV – размер стороны поля зрения, охватывающего объект (мм).

Фокусное расстояние камеры определяется согласно формуле на рисунке 3 [7].

$$f = \frac{S_{sensor}}{2 \cdot \tan\left(\frac{AFOV}{2}\right)} \qquad (3)$$

$$AFOV = 2 \cdot \tan^{-1}\left(\frac{FOV}{2 \cdot WD}\right) \quad (4)$$

Рис. 3. Расчет фокусного расстояния камеры

Где:

WD – рабочее расстояние от объекта до передней линзы объектива (мм),

$AFOV$ – угловое поле зрение в градусах.

Кроме того, способ захвата кадра камерой машинного зрения выбирается в зависимости от того, находится ли изделие в движении или фиксируется на время анализа [6].

Программная архитектура

С учетом особенностей анализируемых объектов и масштабов их крупносерийного производства для автоматизации процесса оценки качества предлагается организация гибкой архитектуры, в которой каждый этап реализован как отдельный программный компонент. Последовательность выполнения контроля включает следующие шаги: позиционирование изделия, захват изображения, предобработку, выявление аномалий, оценку отклонений, принятие решения о годности, формирование протокола и оповещение оператора. Для того, чтобы реализовать такую адаптивную структуру, а также, в отличие от использования нейросетей, сделать ее прозрачной и контролируемой как для разработчика, так и для оператора, в основе функционирования системы лежит применение классических методов машинного обучения. Они позволяют на каждом этапе получать, обрабатывать, передавать и сохранять промежуточные данные.

Также важной отличительной чертой предлагаемой технологии автоматизации является возможность сравнения получаемого изображения детали с эталонами, формируемыми, как минимум, двумя способами – по фотографиям образцов качественных деталей и по двумерным проекциям трёхмерного конструкторского чертежа деталей в виде 3D-моделей. Комплексирование этих данных позволяет повысить точность представления эталона в реальных условиях, реализовать задачу автоматического совмещения проекций эталонов и контролируемой детали, а также пополнения архива деталей с учётом изменения технологического процесса.

Таким образом, ключевыми компонентами программного комплекса являются: подсистема захвата изображения (ПЗИ), модуль добавления эталона (ДЭ), модуль удаления фона (УФ), модуль поиска аномалий (ПА), модуль интерфейса пользователя (ИП).

Перспективы развития

Разработанный прототип реализации описанной технологии подтвердил ее работоспособность. Эксперимент определил факторы, влияющие на условия съемки. При этом для других типов электротехнических изделий потребуется проведение дополнительных экспериментов по подбору параметров, обеспечивающих качественные входные данные.

Дальнейшее развитие системы связано с уточнением конфигурации установки для проведения анализа в условиях конкретного производства, для которого будет запланировано внедрение. Также имеется возможность масштабировать модули архитектуры и внести необходимые программные и аппаратные дополнения для организации многопользовательского доступа. Распараллеливание процессов обработки

изображений позволит повысить скорость работы комплекса для его интеграции в высоконагруженные производственные линии.

В результате, предлагаемый подход может быть положен в основу построения адаптируемой системы автоматического визуального контроля качества электротехнических изделий, учитывающей разнообразие исследуемых объектов и специфику производственных условий.

Список литературы

1. Горелов А.А. Техническое зрение в автоматизированных системах управления. — М.: Радио и связь, 2002. — 272 с.
2. Селиванов А.А., Петров К.И. Автоматизация контроля качества изделий на основе методов ма-шинного зрения // Известия ЮФУ. Технические науки. — 2020. — № 3. — С. 85–92.
3. NI Vision Systems // National Instruments. URL: https://www.ni.com/en-us/shop/vision.html (дата обращения: 04.05.2025).
4. EyeVision Machine Vision Software // EVT Eye Vision Technology. URL: https://www.evt-web.com/en/ (дата обращения: 01.05.2025).
5. Vision Systems – IV Series // Keyence Corporation. URL: https://www.keyence.com/products/vision/vision-sensor/iv-series/ (дата обращения: 05.05.2025).
6. Выбор камеры // Vision Machines. URL: https://visionmachines.ru/articles/s_chego_nachat/vybor_kamery/ (дата обращения: 01.05.2025).
7. Edmund Optics. Imaging System Parameter Calculator // Edmund Optics. URL: https://www.edmundoptics.com/knowledge-center/tech-tools/imaging-system-parameter-calculator/ (дата обращения: 07.05.2025).

ПОСТАНОВКА ПРАКТИКИ ПО РАЗРАБОТКЕ ВЕБ-СИСТЕМЫ ДЛЯ МОНИТОРИНГА ОБЪЕКТОВ И УПРАВЛЕНИЯ ИМИ

ШЕВЧЕНКО А. А.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В данной работе рассматривается разработка практики по дисциплине «Информационные технологии и компьютерная безопасность» для магистров первого курса факультета электротехники и автоматики СПбГЭТУ «ЛЭТИ» по разработке и применению веб-систем в области мониторинга и управления техническими объектами.

Ключевые слова: веб-система, управление, мониторинг, модуль ввода-вывода, макет

Техническое обеспечение практики

Поставленная практика базируется на архитектуре клиент-сервер, которая подразумевает специализацию средств компьютерной техники по функциональному назначению. В их составе выделяются серверы (файловые серверы, серверы баз данных, серверы приложений) и рабочие станции пользователей. Рабочие станции ориентированы на работу с компьютерной графикой, серверы – на хранение и обработку данных [1]. В данной практике участвуют веб-сервер и сервер модуля аналогового и дискретного ввода-вывода.

Рис. 1- схема взаимодействия элементов системы

Сервер модуля аналогового и дискретного ввода-вывода обеспечивает соединение и обмен данными с модулем, к вводам и выводам которого подключены макеты. Данные о состояниях макетов содержатся в файлах данных. Вместо макетов, преимущественно для отладки, возможно использование программ-имитаторов, которые также используют файлы данных аналогично модулю. Студентам предлагается, используя готовое веб-приложение, разработать такую систему страниц и скриптов, которая осуществляет чтение данных из файла данных и запись данных в них для мониторинга за объектами и управления ими. Схема взаимодействия элементов системы представлена на рис. 1.

Разработка макетов и сервера модуля аналогового и дискретного ввода-вывода

В качестве макетов используются четыре железнодорожных светофора, которые представляют собой три светодиода, подключённые через резисторы к цифровым выходам, и кнопку, имитирующую проход поезда, подключённую к цифровому входу (рис. 2). Программа-имитатор представляет собой программную замену макету и модулю с его программой, она позволяет устанавливать состояние светофора и отображает состояние кнопки с датой и временем последней смены состояния. Описания железнодорожных светофоров содержатся в файле данных RailLight_N.txt, где N — номер светофора от 1 до 4. Файл данных имеет следующую структуру: в первой строке содержится индекс состояния (0 — выключен, 1 — красный, 2 — зелёный, 3 — жёлтый), во второй строке находится описание состояния кнопки (0 — не нажата, 1 — нажата), в третьей строке расположена дата и время последней смены состояния кнопки.

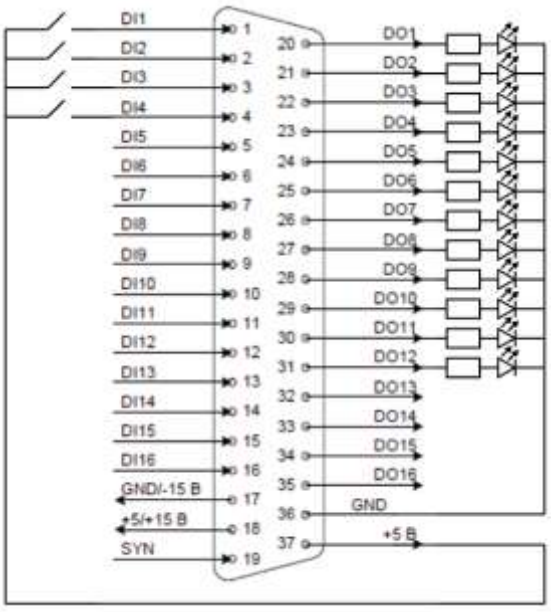


Рис. 2 – электрическая схема макетов

Сервер модуля аналогового и дискретного ввода-вывода Lcard e14-140 обеспечивает соединение и обмен данными с модулем, к вводам и выводам которого подключена периферия в виде макетов. Соединение с модулем, разрешение работы с вводами и выводами и последующий обмен информацией устанавливается через порт usb специальными функциями библиотеки производителя модуля Lcard Lusbapi.lib. Для работы с дискретными вводами и выводами приложение, обеспечивающее работу модуля, формирует или получает 12-ти битное слово, где 2047 тактов соответствуют +5.0 В, 0 - 0 В, –2048 - –5.0 В. Отправка данных на цифровые выходы производится с помощью функции TTL_OUT, которая получает в качестве аргумента слово, а получение данных с входов – с помощью функции TTL_IN, которая записывает слово по переданному указателю в переменную. Приложение регулярно, каждые 500 мс, читает файлы данных объектов и, при несовпадении с текущим состоянием макета устанавливает нужный сигнал на макете или записывает в файл состояние кнопки и время его последнего изменения.

Веб-приложение

При разработке веб-приложения использована библиотека web_lib.h, разработанная на кафедре РАПС СПбГЭТУ “ЛЭТИ”, которая позволяет интерпретировать переданные через URL-запрос параметры. При вызове веб-приложения веб-сервер создаёт список переменных среды, в которые записывается URL-запрос в зависимости от метода передачи данных. При использовании метода GET URL-запрос содержится в переменной среды QUERY_STRING, а при использовании метода POST - передаётся на стандартный ввод процесса, длина запроса содержится в переменной среды CONTENT_LENGTH. При вызове метода Init() осуществляется инициализация запроса - его получение из переменных среды и разбор параметров в структуру pars, которая содержит сведения о подключении. Для сопоставления запроса и функции в приложении используется метод Exec() [1].

Ввиду малого количества часов, отведённого на практику, для студентов подготовлено универсальное веб-приложение, которое осуществляет работу с файлами данных и динамически формирует страницу. В веб-приложении предусмотрена обработка

нескольких запросов: на запись данных в файл, замену строки в файле, чтение данных из файлов, очистку файла и построчное добавление информации в файл. Под запросом понимается идентификатор некоторой функции в приложении, в списке параметров он пишется первым.

Пример выполнение практики

Одной из бригад было дано задание реализовать автоматическое переключение сигналов светофора при проходе поезда по линии, то есть разработать такую систему страниц, скриптов и запросов к веб приложению, чтобы:

1. при проезде поездом первого светофора на нём загорался красный свет;
2. при проезде поездом второго светофора на нём загорался красный свет, на первом – жёлтый;
3. при проезде поездом третьего светофора на нём загорался красный свет, на втором – жёлтый, на первом – зелёным;
4. по истечении заданного времени (10 секунд) на третьем светофоре загорался жёлтый, на втором – зелёный;
5. по истечении заданного времени (5 секунд) на третьем светофоре загорался зелёный.

Страница мониторинга и автоматического управления приведена на рис. 3. Кругами соответствующего цвета обозначены сигналы светофоров, красным прямоугольником обозначено положение поезда на участке пути, также информация о состоянии светофоров, кнопок и последнем времени их срабатывания приведена ниже.

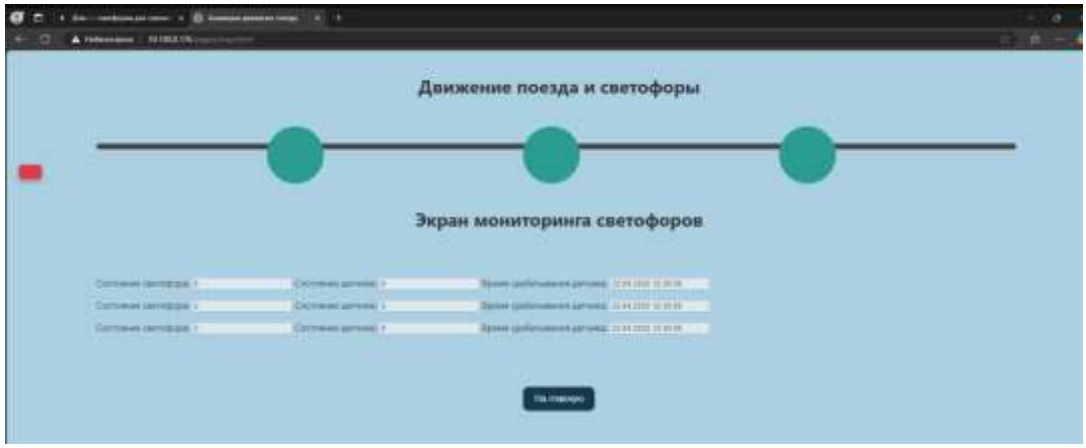


Рис. 3 - страница мониторинга и автоматического управления

Исходя из задания, студентами разработан алгоритм работы скрипта:

1. Поезд не на пути, все светофоры зеленые.
2. При изменении состояния датчика с 0 на 1 на первом светофоре, что равносильно проеду поезда мимо него, цвет первого светофора меняется на красный.
3. При изменении состояния датчика с 1 на 0, поезд едет дальше.
4. При изменении состояния датчика с 0 на 1 на втором светофоре, цвет второго светофора меняется на красный, первого - на желтый.
5. При изменении состояния датчика с 1 на 0, поезд едет дальше
6. При изменении состояния датчика с 0 на 1 на третьем светофоре, цвет третьего светофора меняется на красный, второй светофор становится желтым, первый светофор становится зеленым. Флаг состояния выставляется в 1.

7. Выполняется задержка по флагу – спустя 10 секунд третий светофор становится желтым, а второй зеленым. Флаг выставляется в 2.
8. Выполняется задержка по флагу – спустя 5 секунд третий светофор становится зеленым, флаг сбрасывается.
9. На каждой итерации выполнения асинхронной функции считываются значения из файлов и выводятся в окна на части экрана, предназначенного для мониторинга.

В скрипте были реализованы следующие функции:

- Функция setLight(number, color) выполняет вызов веб-приложения для изменения состояния в файле данных соответствующего светофора
- Функция updateSignals обновляет отображение цвета светофора на веб-странице
- Положение поезда на экране изменяется в зависимости от состояния кнопки (имитация датчика движения) при помощи функции updateTranePos(pos)
- Асинхронная функция poll() выполняет опрос состояния светофоров и датчиков, а также реализует логику их изменения
- Переходы между состояниями и задержки реализованы по условию switch (flag)
- Функция GetPars() запрашивает содержимое текстовых файлов со светофорами и извлекает нужные значения.

Заключение

Архитектура клиент-сервер, лежащая в основе практики, продемонстрировала свою эффективность в организации взаимодействия между различными компонентами системы. Серверная часть успешно обеспечила обмен данными с модулями и обработку запросов от клиентов. Использование файлов данных для хранения состояний макетов и возможность замены реальных макетов на программы-имитаторы значительно упростило процесс отладки и тестирования системы, что особенно важно при обучении студентов. Разработка системы веб-страниц и скриптов для мониторинга и управления объектами позволила студентам получить практический опыт. Таким образом, представленная практика успешно продемонстрировала работу системы мониторинга и управления на основе клиент-серверной архитектуры и обеспечило её понимание студентами.

Список литературы

1. Шевченко А. В. Компьютерные, сетевые и информационные технологии: Учеб. пособие. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2013. 64 с.

СЕКЦИЯ СИСТЕМНЫЙ АНАЛИЗ И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

АНАЛИЗ ЭФФЕКТИВНОСТИ ВЫДЕЛЕНИЯ ЭЛЕМЕНТОВ БИЗНЕС-ПРОЦЕССОВ С ПОМОЩЬЮ ЯЗЫКОВОЙ МОДЕЛИ ChatGPT

А.С. Буравлев, К.А. Суслова, А.С. Прохоров

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И.Ульянова (Ленина)

Аннотация. В данной статье целью исследования является анализ возможностей повышения эффективности трудовой деятельности бизнес-аналитиков в части реконструкции бизнес-процессов с помощью языковой модели ChatGPT 4.0. Научная новизна работы заключается в получении ранее неизвестных результатов эффективности ChatGPT для реконструкции отдельных проекций бизнес-процессов(поток управления, данные, ресурсы, операции) на основании изучения результатов реконструкции 54 коротких пользовательских сценариев с последующим сравнительным анализом с существующим подходом реконструкции процессов на основе правил грамматики зависимостей. Практическая значимость исследования обусловлена возможностью использования полученных данных для уточнения схем обработки пользовательских историй. В работе решаются следующие задачи: анализ архитектуры и возможностей модели ChatGPT 4.0 в части обработки естественного языка, разработка методики оценки качества реконструкции бизнес-процессов, экспериментальная оценка качества реконструкции, получение статистических оценок, сравнительный анализ с существующим подходом на основе правил грамматики зависимостей. Для достижения поставленных задач в работе используется аппарат статистической обработки данных, экспертного анализа, прикладной лингвистики и нейронных сетей.

Ключевые слова: бизнес-процесс, обработка естественного языка, шлюзы, нейронные сети, модель трансформер, языковые модели, ChatGPT

Введение.

Текущий курс развития информационных систем [1], [2] характеризуется переходом от данных к процессам. В центре автоматизации теперь находятся не структуры и модели данных, а бизнес-процессы как основа управления. Концентрация на целях предприятия и способах их достижения в рамках данного подхода позволяет учитывать изменения в автоматизируемых системах, а фиксированные характеристики объектов отодвинуть на второй план.

В производственной среде бизнес-процессы создаются бизнес-аналитиками на основе анализа нормативной документации и опросов сотрудников. Чаще всего они представлены в виде некоторого набора сценариев (пользовательских историй), формализация которых осуществляется с использованием языков моделирования, таких как BPMN или IDEF. Процессы являются гибкими и подлежат доработке в случае изменения требований автоматизации. Однако формализация пользовательских историй зачастую требует высокой квалификации от бизнес-аналитика, поскольку представляет собой трудоемкую задачу. Для ее решения существует множество подходов обработки естественного языка с различной степенью автоматизации, основанных как на обнаружении в тексте «характерных шаблонов» элементов, так и на формировании некоторого промежуточного представления [2].

На этом фоне стремительное развитие методов обработки естественного языка (NLP) открывает новые возможности для автоматизации подобных задач. Появление архитектуры Transformer, представленной в работе "Attention is All You Need" [3], стало ключевым

этапом в развитии NLP, а её внедрение в модели нового поколения, такие как ChatGPT [4], кардинально изменило подход к обработке текстовой информации. ChatGPT продемонстрировал возможность преобразования неструктурированного текста в структурированные данные, что делает его полезным инструментом для анализа пользовательских историй и моделирования бизнес-процессов.

Методология создания процессов из текстовых сценариев включает несколько этапов [2]:

1. **Определение границ процесса** — выявление событий, запускающих процесс, и определение его возможных результатов.
2. **Определение действий** — установление элементарных операций, задействованных в процессе.
3. **Определение ресурсов и порядка передачи управления** — идентификация участников процесса (компаний, отделов и других зон ответственности), определение ролей и последовательности смены исполнителей.
4. **Определение потока управления** — установление порядка выполнения действий, включая параллельные, условные и циклические сценарии.
5. **Добавление дополнительных элементов** — расширение модели за счет включения объектов, с которыми связаны элементарные операции (например, изделия, документы и т. д.).

В данной работе при условии, что границы процесса являются заданными пользовательской историей (по этапу 1) был поставлен эксперимент по оценке реконструкции модели процесса (этапы 2-5) с помощью языковой модели ChatGPT 4.0 на основании 54 коротких пользовательских сценариев (до 300 слов). Произведён сравнительный анализ качества реконструкции этапов 2-5 с результатами применения подхода на основе правил выделения элементов из дерева, построенного на основе формируемой автоматически грамматики зависимостей [2].

Описание языковой модели ChatGPT

ChatGPT – является языковой моделью, разработанной компанией OpenAI. Механизм работы модели основан на вероятности появления следующего слова в заданной последовательности и генерации текста на основе входных данных. Языковые модели, такие как ChatGPT, разработаны на основе архитектуры трансформера, которая позволяет эффективно анализировать и генерировать текст. Поскольку основной задачей большинства языковых моделей является предсказание следующего слова (токена), в них используется только часть модели трансформера, а именно, декодер.

Attention — это ключевой компонент архитектуры трансформера [3]. Он позволяет модели фокусироваться на наиболее значимых частях текста. Attention вычисляет вес для каждого токена относительно всех остальных. Этот вес показывает, насколько сильно один токен влияет на другой. В источнике [5], подтверждается, что отличие от стандартной архитектуры трансформера, в GPT-3 используется расширение базового механизма внимания, под названием Multi-Head Attention. Так как официально компания OpenAI не опубликовывала деталей архитектуры GPT-4, можно предположить, что данная модель унаследовала данное расширение механизма. Согласно [3], Multi-Head Attention позволяет модели совместно обрабатывать информацию из разных подпространств представлений, что критически важно для разнородных зависимостей в данных.

На основании представленного расширения механизма внимания, авторами выдвинута гипотеза, что в контексте этапов реконструкции модели процесса из пользовательских сценариев механизм Multi-Head Attention (предположительно используемый) в GPT-4 может улучшить результат за счет параллельного анализа разноуровневых зависимостей (каждый набор “голов” фокусируются на разного рода задачах, одни отвечают за выявление порядка этапов процесса, другие за определение ролей акторов, а третьи за логические ветвления и т.д.) и адаптации к вариативности формулировок (Multi-Head Attention агрегирует признаки из разных подпространств, чтобы сопоставить синонимичные действия). Наиболее значительным преимуществом механизма Multi-Head Attention является его способность к обучению, что позволяет GPT-4 адаптироваться к специфике бизнес-процессов без ручной настройки правил

Результаты эксперимента.

Целью исследования является оценка качества реконструкции бизнес-процессов с помощью языковой модели ChatGPT.

Эксперимент был разделён на 4 этапа:

1. Получение результатов обработки пользовательских сценариев языковой моделью ChatGPT.
2. Экспериментальная оценка качества реконструкции.
3. Расчет статистических показателей.
4. Сравнительный анализ качества реконструкции бизнес-процессов с подходом, использованным в источнике [2].

Получение результатов обработки пользовательских сценариев языковой моделью ChatGPT. В соответствии с приведенной во «Введении» методикой моделирования процессов языковой модели ChatGpt 4 была поставлена задача по реконструкции элементов бизнес-процессов, выделенных в этапах 2-5. Считая границы процесса заданными, через API нейросетевой модели было отправлено 54 запроса, каждый из которых содержал задачу по выделению действий; И/ИЛИ шлюзов; субъектов и объектов; пулов и дорожек. В экспериментальном исследовании использовалась выборка из 54 коротких (до 300 слов) текстовых сценариев.

Экспериментальная оценка качества реконструкции. На этапе оценки качества реконструкции авторами был осуществлен анализ качества реконструкции бизнес-процессов с помощью модели ChatGPT. Оценка обработки нейросетевой моделью набора пользовательских сценариев производилась на основании 4-х критериев, определяющих число корректно распознанных экземпляров по процентной шкале от 0 до 100%.

Расчет статистических показателей. По полученным данным оценок, в каждом критерии были выделены минимальное и максимальное значение (min и max), получены такие величины как размах выборки (R), количество и длина интервалов (h). На основании середин интервалов и абсолютных частот (количества значений m_i, попавших в определенный интервал) были построены полигоны абсолютных частот. Ключевыми рассчитанными статистическим оценками стали: дисперсия (D), среднеквадратичное отклонение (σ), среднее арифметическое (μ), мода (M_o) и медина (M_e). Все расчеты были производились полуавтоматически, с помощью средств, предоставляемых пакетом Excel.

Сравнительный анализ качества реконструкции бизнес-процессов с подходом на основании метода обработки естественного языка. На финальном этапе эксперимента была произведена сравнительная оценка обработанных статистических данных на

основании применения модели ChatGPT 4o и результатами применения подхода на основе правил грамматики зависимостей из источника [2].

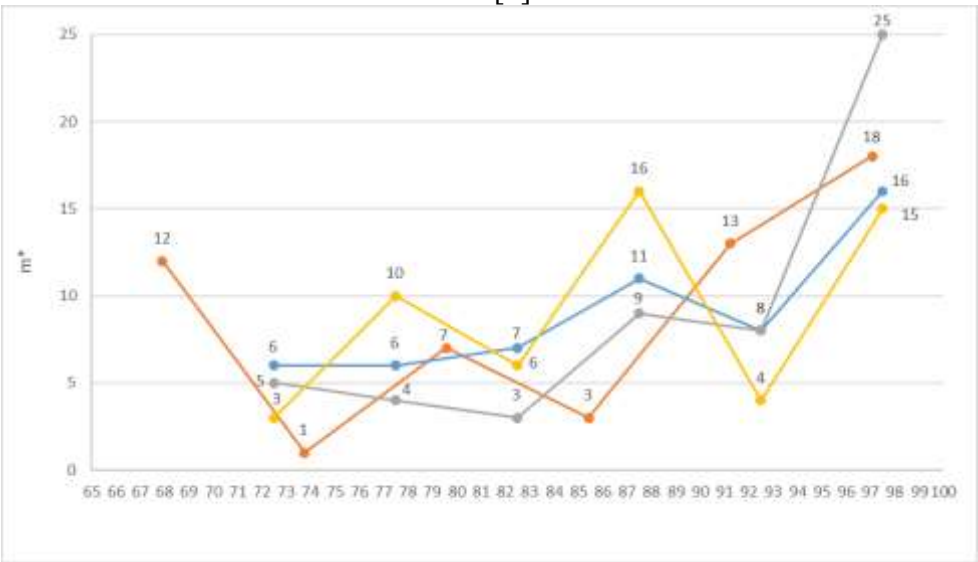


Рис.1. Полигоны частот по 4-м критериям

На рисунке 1 оранжевым цветом обозначена линия шлюзов, синим – объекты, серым – ресурсы, желтым – действия. Ось X – процентная шкала, ось Y – количество оценок.

Таблица 1

Статистические оценки

	Оценка действий	**Оценка ресурсов**	**Оценка объектов**	**Оценка шлюзов**
$\boldsymbol{\mu}$	89,7	92,6	90,1	86,4
$\boldsymbol{D}$	71,1	87,5	80,1	148,5
$\boldsymbol{\sigma}$	8,4	9,4	9	12,1
$\boldsymbol{M_e}$	90	95	90	90
$\boldsymbol{M_o}$	90	100	100	100

Таблица 2

Сравнительные характеристики подходов

	Оценка действий	**Оценка ресурсов**	**Оценка объектов**	**Оценка шлюзов**
Подход на основе модели ChatGPT	89,7	92,6	90,1	86,4
Подход на основе правил грамматики зависимостей [2]	89	90	78	50

Выводы. Обсуждения.

Полученные статистические оценки среднего в рамках эксперимента на этапе обработки данных оценок на основании применения языковой модели ChatGPT отражают высокий потенциал модели в ее применении для автоматизации всех этапов реконструкции бизнес-процесса (действий, ресурсов, объектов и шлюзов). В отличие метода [2] подход обработки естественного на основе модели ChatGPT при примерно равных значениях качества реконструкции действий и ресурсов продемонстрировал значительные преимущества в части выделения шлюзов (на 36,4%) и объектов (на 12,1%).

Дальнейшие исследования авторы связывают с расширением выборки пользовательских сценариев для получения более значимых статистических оценок в части выявления И/ИЛИ шлюзов. Таким образом, планируется детально исследовать качество определения шлюзов, а также более точно охватить каждый этапов реконструкции бизнес-процессов.

Список литературы

1. Ван дер Аальст В., Ван Хи К. Управление рабочими процессами: модели, методы и системы / пер. с англ. В. А. Башкин, И. А. Ломазова; под ред. И. А. Ломазовой. — М.: Логос, 2007. — 526 с.
2. Буравлев А. С., Демидова Д. Е., Васильев Н. В. Подход к выявлению элементов моделей бизнес-процессов из текстов на естественном
3. Vaswani A., Shazeer N., Parmar N., et al. Attention is all you need // Advances in Neural Information Processing Systems. — 2017. — Vol. 30. — P. 5998–6008
4. Документация ChatGPT [Электронный ресурс]. — Режим доступа: https://openai.com/index/chatgpt/ (дата обращения: 20.04.2025).
5. Brown T., Mann B., Ryder N., et al. Language models are few-shot learners // Advances in Neural Information Processing Systems. — 2020. — Vol. 33. — P. 1877–1901

ИДЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЕЙ ПО ГЕОМЕТРИИ РУКИ: МЕТОДЫ И ПЕРСПЕКТИВЫ

БУШУЕВ Д.И., ШКЛЯР Е.В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Аннотация. В статье рассмотрены современные подходы к биометрической идентификации с использованием геометрии руки. Описаны этапы сбора и предобработки данных, методы выделения характерных признаков и алгоритмы сопоставления. Проанализированы преимущества и ограничения метода по сравнению с другими биометрическими признаками. Показано, что геометрия руки обеспечивает высокую надёжность и скорость распознавания при относительной простоте аппаратной реализации.

Ключевые слова: биометрия, геометрия руки, идентификация, биометрический шаблон, распознавание

Введение

Биометрическая идентификация по геометрии руки зарекомендовала себя как надёжный (с приемлемой точностью) и экономичный (с не сложной аппаратурой) метод аутентификации [8]. В отличие от систем распознавания отпечатков пальцев или радужки глаза, геометрия руки менее чувствительна к загрязнениям сенсора и требует более простого оборудования, что делает её привлекательной для массовых приложений контроля доступа и учёта рабочего времени [1], [2]. Каждый человек, согласно данным исследований [1], [4], демонстрирует индивидуальные особенности в структуре кисти, что позволяет использовать геометрию руки для задач биометрической идентификации. Однако изменения кисти с возрастом все еще имеют свой вес, поэтому требуют актуализации шаблона.

Описанные ниже этапы обработки изображений соответствуют общей схеме построения биометрических систем, приведенной в ГОСТ Р 54412–2019

Сбор данных и предобработка

Для съёмки контура руки чаще всего используется плоский оптический сенсор с разрешением ≥300 dpi, что обеспечивает погрешность измерений менее 0,5 мм [3]. Важным условием является равномерная инцидентная подсветка — LED-панель под матовым рассеивателем — и отсутствие бликов на стекле.

Обработка осуществляется с помощью алгоритмов сегментации:

– *Пороговая фильтрация*. Глобальный или адаптивный порог (например, метод Отсу) разделяет пиксели на фон и объект [3], [4];

– *Морфологическая очистка*. Операции эрозии и дилатации удаляют мелкие шумовые артефакты и заполняют «дыры» в силуэте кисти [3], [4];

– *Поиск ключевых точек*. Алгоритм Хафа или метод поиска контуров находит экстремумы по вертикали — кончики пальцев и основание ладони [9].

Полученный бинарный контур приводится к единому масштабу: расстояние между точками основания среднего и безымянного пальцев фиксируется (например, 100 мм). Угол поворота ладони компенсируется поворотом изображения так, чтобы ось ладони оказалась вертикальной.

Извлечение признаков

К линейным размерам относятся длины пальцев $L_i, (i = 1 \dots 5)$ измеряемые от основания до кончика, и ширина ладони в трех точках: у основания пальцев, в середине ладони и у запястья (W_j).

Угловые показатели включают углы раскрытия между пальцами (φ_{12} между осями пальцев *I–II* и φ_{23} между осями пальцев *II–III*).

Контурные характеристики отражают радиусы кривизны суставов (R_k) и асимметрию ладони, вычисляемую по формуле:

$$\mathrm{A} = \frac{W_1 - W_2}{W_1 + W_2}$$

где W_1и W_2 – ширины в двух соседних сечениях.

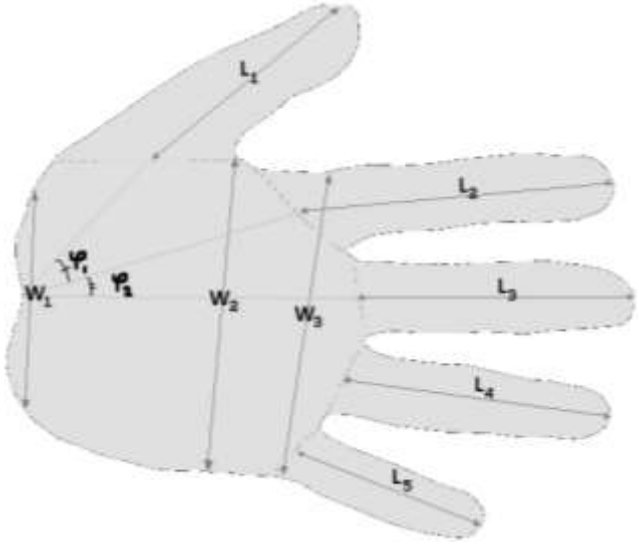


Рис. 1. Схема руки с условными обозначениями

Методы сопоставления

Для обеспечения работы подсистемы сопоставления в системах идентификации по геометрии руки применяются различные методы классификации.

Метод k-ближайших соседей (k-NN) осуществляет сравнение признаковых векторов с использованием евклидова расстояния. Параметр k подбирается на этапе кросс-валидации. Преимуществами метода являются простота реализации и отсутствие фазы явного обучения [7], [8].

В исследовании [4] в качестве метода сравнения использовались гауссовские смешанные модели (GMM). Их особенность в том, что они представляют распределение признаков как смесь нормальных распределений, параметры которых оцениваются с помощью алгоритма ожидания-максимизации (EM).

Применение нейронных сетей (MLP, CNN) позволяет повысить устойчивость распознавания за счёт автоматического извлечения сложных признаков из изображения [5].

Экспериментальная оценка

Для оценки практической эффективности алгоритмов анализа геометрии руки в литературе используют выборки. В исследованиях [7], [8] участвовали десятки человек, в то время как в работах [4] рассматривались выборки в несколько сотен участников. Как правило, каждому испытуемому снимают несколько (3–10) изображений кисти с небольшими вариациями положения и освещения, после чего данные разбивают на тренировочную и тестовую части (например, в соотношении 70%/30%) и применяют кросс-валидацию для повышения надёжности результатов.

В разных исследованиях сообщается о сходных диапазонах точности распознавания:

– Raul Sánchez-Reillo и соавторы [8] при использовании линейных и контурных признаков достигли точности идентификации порядка 95%;

– Mays & Taylor [7] сообщили об уровне распознавания около 93%, применяя k-NN на векторных описаниях геометрии;

– Артамонова и Барсуков [4] при сочетании гауссовских смесей и отбора наиболее информативных признаков показали точность свыше 96%.

Время обработки одного шаблона (включая сегментацию, нормализацию и классификацию) в современных реализациях обычно не превышает 50–100 мс на процессорах уровня Intel Core i5/i7 без специализированных ускорителей, что удовлетворяет требованиям систем контроля доступа в реальном времени.

Основные выводы:

– Границы точности (90–98%) зависят от выбора признаков и алгоритмов классификации;

– Больше всего выигрыша даёт мультимодальная схема, когда геометрия руки дополняется, например, отпечатком пальца или венозным рисунком;

– Важным остаётся корректное положение кисти и качество сегментации: смещение более 5 мм или нерезкий контур может снижать точность на несколько процентов.

Достоинства и ограничения метода идентификации пользователей по геометрии руки

Преимущества:

– Компактность шаблона (≤20 признаков) и быстрота сравнения;

– Низкая стоимость сенсора — стандартный плоский оптический сканер разрешением 300 dpi (без специализированной оптики) — и устойчивость к загрязнениям за счёт применения светорассеивающей поверхности.;

– Умеренная вариативность признаков со временем (рост не более 2–3 % в год по данным [8], на основе многолетних наблюдений динамики биометрических параметров рук).

Ограничения:

– Снижение точности при очень похожих строениях кистей, таких как у однояйцевых близнецов, зафиксировано в исследовании [8];

– Необходимость строгой ориентации и положения руки на сенсоре: в работах [7], [8] показано, что смещение более 5 мм приводит к снижению точности идентификации на 2–5%.;

– Возможность ошибки сегментации при наличии браслетов, часов или перстней.

Перспективы развития

– *Мультимодальная биометрия*. Объединение геометрии руки с венозным рисунком ладони и отпечатком пальцев для повышения надежности и защиты от подделок [6];

– *Мобильные приложения*. Применение камер смартфонов с алгоритмами коррекции позы и освещения для дистанционной аутентификации;

– *Глубокое обучение*. Разработка специализированных архитектур CNN с учётом особенностей ладонных изображений и малых обучающих выборок;

– *Интернет вещей (IoT):* встраивание биометрических сенсоров (сканеров кисти) в бытовые устройства — например, в системы умного дома для индивидуальной настройки доступа или активации профилей пользователей.

Заключение

Геометрия руки остаётся одним из наиболее сбалансированных по «цена–качество» биометрических методов: он не требует дорогостоящей оптики и при этом обеспечивает надёжное распознавание в подавляющем большинстве практических сценариев. Вместе с тем реализация системы на практике требует тщательной отладки аппаратной части — сенсора и подсветки — а также разработки удобного интерфейса для пользователя, гарантирующего правильное положение кисти.

Основной вызов для дальнейших разработок — повышение устойчивости к вариациям положения и объективным условиям (освещение, аксессуары на руке) без значительного усложнения алгоритмов. Решением может стать гибридный подход: комбинировать геометрия руки с дополнительными признаками (венозный рисунок, отпечатки пальцев) или внедрять адаптивные нейронные сети, способные автоматически корректировать ошибки сегментации.

Наконец, важным направлением остаётся интеграция таких систем в мобильные и IoT-устройства. Упрощение процесса съёмки с помощью камер смартфонов и перенос части вычислений в облако позволит сделать биометрию кисти доступной везде, где сегодня востребована быстрая и безопасная идентификация.

Список литературы

1. ГОСТ Р 54412. Информационные технологии. Биометрия. Общие положения и примеры применения;
2. ISO/IEC 24741. Information technology. Biometrics. Overview and application;
3. ГОСТ ISO/IEC 2382-37. Информационные технологии. Словарь. Биометрия;

4. Артамонова Я. В., Барсуков С. С., Суханова А. А. Биометрические технологии: отношение населения и преодоление рисков // Государственное и муниципальное управление. Ученые записки. 2024. № 1. С. 231–236. https://doi.org/10.22394/2079-1690-2024-1-1-231-236. EDN ABUSEC;
5. Храмов М.А., Корнев Л.В., Шабля В.О. Феноменологический анализ существующих методов аутентификации // Научный результат. Информационные технологии. – Т.9, №1, 2024. – С. 29-36. DOI: 10.18413/2518-1092-2024-9-1-0-4;
6. Чекмарев Д.Ю., Борзенкова С.Ю. Проблемы идентификации пользователей в информационных системах // Известия ТулГУ. Технические науки. 2023. Вып. 3;
7. Mays M. Tayler, Dr. Loay E. George. A digital signature system based on hand geometry // Wasit Journal of Computer and Mathematic Science. Vol. 1 No. 1 2022;
8. Raul Snchez-Reillo, Carmen Snchez-Avila, Ana Gonzalez-Marcos. Biometric Identification through Hand Geometry Measurements // IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 22 No. 10 October 2000;
9. Rafael C. Gonzalez, Richard E. Woods. Digital Image Processing Fourth Edition.

МЕТОДЫ ОПРЕДЕЛЕНИЯ ВИТАЛЬНОСТИ В БИОМЕТРИЧЕСКИХ СИСТЕМАХ

Иконников С.Н., Шкляр Е.В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В. И. Ульянова (Ленина)

Аннотация. Исследование посвящено методам определения витальности (liveness detection), которые играют ключевую роль в защите от атак на биометрическое предъявление (АБП). Особое внимание уделяется пассивным методам определения витальности, использующим нейросетевые подходы для распознавания живых субъектов. Основными аргументами исследования являются обзор современных методов, таких как ансамбли нейросетевых моделей, использующих различные методы обработки изображений, а также анализ существующих ограничений и перспектив развития этих технологий.

Ключевые слова: биометрия, определение витальности (liveness detection), обнаружение атак на биометрическое предъявление.

Введение

Технологии биометрической аутентификации становятся более распространёнными: с помощью биометрии можно разблокировать телефон, попасть в общежитие или на работу. По данным Единой Биометрической Системы, в России уже сейчас можно использовать биометрию для беспарольной аутентификации на портале «Госуслуги», в банковских приложениях, для взаимодействия с ФНС и МФЦ в некоторых регионах, выпуска квалифицированной ЭЦП, для оформления карты болельщика. Биометрия становится всё более распространённой для оплаты покупок: повсюду можно встретить терминалы позволяющие расплатиться при помощи лица в магазинах и общественном транспорте [1].

Некоторые из этих систем функционируют без постоянного надзора, что делает их уязвимыми к атакам на биометрическое предъявление: злоумышленник может использовать различные методы для получения доступа к системе путём выдачи себя за легитимного пользователя. В связи с этим становится актуальной проблема обнаружения подобных атак. В этой работе мы сосредоточимся на определении витальности, методе ОАБП для определения того, что данные получены от живого субъекта.

Цели исследования

Целью исследования является систематизация современных методов определения витальности в биометрических системах, анализ подходов к распознаванию атак на биометрическое предъявление, а также выявление текущих тенденций, ограничений существующих решений и перспектив развития технологии.

Термины и определения

ГОСТ Р 58624.1–2019 [2] вводит определения, связанные с биометрическими системами и их защитой. Используемые в статье термины приведены ниже.

Витальность – свойство или состояние живого индивида, подтверждаемое анатомическими характеристиками, непроизвольными реакциями, физиологическими функциями, произвольными реакциями или поведенческими характеристиками индивида.

Атака на биометрическое предъявление (АБП) – биометрическое предъявление подсистеме сбора биометрических данных с целью вмешательства в работу биометрической системы.

Обнаружение атаки на биометрическое предъявление (ОАБП) – автоматическое обнаружение атаки на биометрическое предъявление.

Инструмент атаки на биометрическое предъявление (ИАБП) – биометрическая характеристика или объект, используемые для атаки на биометрическое предъявление.

Обнаружение витальности – измерение и анализ анатомических характеристик, непроизвольных или произвольных реакций индивида с целью определения того, что биометрический образец получен от живого индивида.

Описание атаки

Рассматривается атака с помощью искусственно созданных ИАБП, включающих в себя распечатанные фотографии, фотографии и видеозаписи, предъявляемые системе на экране мобильного устройства, маски, имитирующие лицо легитимного пользователя (бумажные, тканевые, силиконовые, с вырезами для глаз, носа и т.д.).

Метрики оценки эффективности

Для сравнения методов определения витальности будут использоваться метрики:

1. APCER (attack presentation classification error rate), вероятность ошибки классификации предъявления при атаке – доля предъявлений при атаке одним видом инструментов атаки на биометрическое предъявление, которые некорректно классифицируют как подлинные биометрические предъявления в конкретном сценарии [3];
2. BPCER (bona fide presentation classification error rate), вероятность ошибки классификации подлинных биометрических предъявлений – доля подлинных биометрических предъявлений, которые некорректно классифицируют как предъявления атаки в конкретном сценарии [3];
3. ACER (average classification error rate), средняя вероятность ошибки классификации – комбинированный показатель, рассчитываемый как среднее арифметическое значений APCER и BPCER, показатель средней ошибки системы.

Анализ предложенных методов

Для анализа методом сплошной выборки выбирались статьи с arXiv, ScienceDirect и LivDet по ключевым словам «face liveness detection», «face spoofing detection», «face presentation attack» за период с 2021 по 2025 год.

Рассмотренные работы сосредоточены на пассивном определении витальности, для принятия решений используются различные нейросети, обученные на наборах данных 3D MAD, RECOD-MPAD, OLU-NPU, SiW и ряде других, в том числе собственных наборов. Наборы содержали различное количество субъектов и состояли из изображений и видеозаписей, содержащих как нормальные биометрические предъявления, так и АБП с использованием различных типов ИАБП: фотографии, 2D- и 3D-маски, изображение с экранов устройств, силиконовые маски. Для удобства, основные метрики для рассмотреных ниже методов сведены в таблицу 1.

В ходе соревнования LivDet – Face 2021 наиболее эффективно показали себя подходы [4]: ансамбль из трёх моделей – DeepPixBis для анализа каждого пикселя входных данных, ResNetXt, обученный на наборе данных ImageNet, выдающий бинарный результат и легковесная модель, анализирующая входные данные в частотном диапазоне (APCER = 17,61%, BPCER = 15,33%, ACER = 16,47%). Модели достойно показали себя против высококлассных ИАБП, таких как 3D маски. Второй подход использовал модель с 12 слоями, обученную с нуля на публичных наборах данных, с применением центральной разностной свертки и нескольких модулей внимания для улучшения производительности (APCER = 13,33%, BPCER = 24,08%, ACER = 18,71%). Несмотря на уступающие показатели, модель лучше предыдущей справилась с простыми ИАБП – распечатанными изображениями, экранами устройств и простыми 2D- и 3D-масками.

В аналогичном соревновании 2024 года [5] команды продемонстрировали более эффективные результаты, а среди самых эффективных подходов снова можно увидеть ансамбль моделей, а именно сочетание EfficientNet с Vision Transformers, объединяющих высокую пропускную способность и возможность выделения наиболее важных для задачи зон входного изображения (APCER = 0,46%, BPCER = 10,60%, ACER = 5,53%), модели хорошо показали себя на тканевых и 3D масках. Второй же подход использовал автокодировщик на основе VGG19, а при обучении использовались только подлинные биометрические предъявления (APCER = 0,88%, BPCER = 18,93%, ACER = 9,91%), модель хорошо показала себя против таких ИАБП как фотографии, маски и проекции.

В статье [6] авторы предложили метод LwFLNeT – лёгкую двухпоточную архитектуру глубоких сверточных нейронных сетей (APCER = 0,2%, BPCER = 0,4%, ACER = 0,3%). Одна из ветвей модели отвечает за низкоуровневые признаки, вторая же работает с более комплексными структурами, модель хорошо показала себя при работе с 2D- 3D-масками, обойдя по показателям модели VGG-16 и RESNET-50.

Авторы статьи [7] предлагают очередную ансамблевую модель, состоящую из нескольких легковесных свёрточных моделей, входные данные передаются на каждую модель и на основе решений всех моделей выносится итоговое (APCER = 0%, BPCER = 5,355%, ACER = 2,67%), модель показала хорошие результаты при работе с различными типами ИАБП, включая 2D- и 3D-маски.

Таблица 1

Метод	APCER, %	BPCER, %	ACER, %
LivDet-2021-1 [4]	17,61	15,33	16,47
LivDet-2021-2 [4]	13,33	24,08	18,71
LivDet-2024-1 [5]	0,46	10,60	5,53
LivDet-2024-2 [5]	0,88	18,93	9,91
LwFLNeT [6]	0,2	0,4	0,3
Ensemble Model [7]	0	5,355	2,67

Ограничения

Анализ источников показывает, что наибольшей эффективностью в определении АБП обладают ансамбли моделей, каждая из которых отвечает за собственную задачу или признак при анализе входного изображения для обнаружения витальности. Каждая из моделей должна быть обучена, что может занимать много времени, а работа может требовать значительных ресурсов, поэтому в исследованных работах стремились предложить легковесные модели [6][7].

Ещё одной потенциальной проблемой являются наборы данных, на которых обучаются модели: авторы [6] заявляют, что хотя 3D MAD, NUAA и Replay Attack покрывают разные типы атак, у этих наборов есть ограничения. Например, мало вариативности по позам, фону, условиям окружающей среды. Наборы созданы в контролируемых условиях и не полностью отражают реальные условия.

Эти факторы могут сильно усложнять использование рассмотренных методов в реальных условиях, мобильные устройства, например, могут быть требовательны к энергопотреблению и вычислительным мощностям.

Заключение

В данной работе была проведена систематизация современных методов обнаружения витальности в биометрических системах, проанализированы подходы к распознаванию атак на биометрическое предъявление, а также выявлены ключевые тенденции, ограничения и перспективы развития технологии.

Исследование показало, что наиболее эффективные решения основаны на ансамблях нейросетевых моделей, сочетающих различные методы анализа изображений: частотные признаки, внимание к ключевым областям, многоуровневую обработку данных. Это позволяет достичь высоких показателей точности (APCER < 1%, BPCER < 10% в лучших моделях), особенно против сложных атак, таких как 3D-маски и силиконовые маски. Однако такие системы требуют значительных вычислительных ресурсов и сложны в обучении, что ограничивает их применение в реальных условиях.

Несмотря на высокие показатели рассмотренных методов, остаётся ряд проблем, которые требуют решений. Перспективными направлениями в данной области являются: разработка наборов данных для более точного обучения моделей и возможности отслеживать большее количество атак, оптимизация ансамблевых решений для лучшей доступности в реальных условиях, использование методов запрос-ответ для усложнения атак.

Список литературы

1. Единая биометрическая система URL: https://ebs.ru/ (дата обращения: 23.04.2025).
2. ГОСТ Р 58624.1–2019 (ИСО/МЭК30107 – 1:2016) Информационные технологии БИОМЕТРИЯ Обнаружение атаки на биометрическое предъявление Часть 1 Структура.
3. ГОСТ Р 58624.1–2019 (ИСО/МЭК30107 – 1:2016) Информационные технологии БИОМЕТРИЯ Обнаружение атаки на биометрическое предъявление Часть 3 Испытания и протоколы испытаний.
4. Sandip Purnapatra, Nic Smalt, Keivan Bahmani, Priyanka Das, David Yambay, Amir Mohammadi, Anjith George, Thirimachos Bourlai, Sebastien Marcel, Stephanie Schuckers. Face Liveness Detection Competition (LivDet-Face) – 2021.
5. Lambert Igene, Afzal Hossain, Mohammad Zahir Uddin Chowdhury, Humaira Rezaie, Ayden Rollins, Jesse Dykes, Rahul Vijaykumar, Alain Komaty, Sébastien Marcel, Stephanie Schuckers. Face Liveness Detection Competition (LivDet-Face) – 2024.

6. Swapnil R. Shinde, Anupkumar M. Bongale, Deepak Dharrao, Sudeep D. Thepade. An enhanced light weight face liveness detection method using deep convolutional neural network.
7. Avinash Patel, Mrinal Haloi, Asif Salim, Shashank Shekhar. An Ensemble Model for Face Liveness Detection.

ПРОГРАММНЫЙ МОДУЛЬ ОПРЕДЕЛЕНИЯ ОШИБОК ПРИ РАЗРАБОТКЕ ПОЛЬЗОВАТЕЛЬСКИХ ИНТЕРФЕЙСОВ

КАТРУЩЕНКО О.Д.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В данной статье описывается процесс создания программного модуля, определяющего ошибки в дизайне пользовательского интерфейса веб-сайтов по двум группам критериев: доступность веб-контента и использование и сочетания цветов. Список критериев был сформулирован на основе изученного теоретического материала, содержащего требования и рекомендации по проектированию эффективного пользовательского интерфейса. Для демонстрации работы модуля создан веб-сайт, демонстрирующий функциональные возможности программного модуля.

Ключевые слова: пользовательский интерфейс, UX/UI-проектирование, доступность веб-контента, веб-сайт

Введение

В связи с активным внедрением цифровых технологий во все сферы жизни, число веб-сайтов и их востребованность растет, следовательно, появляются определенные требования к их оформлению. Чрезвычайно важно организовать эффективное обучение специалистов, принимающих участие в создании внешнего вида веб-сайтов, так как некачественный интерфейс веб-сайта может привести к ухудшению пользовательского опыта, потере клиентов. Созданный в рамках данной работы программный модуль призван помочь начинающим специалистам увидеть свои ошибки, узнать актуальные принципы проектирования пользовательских интерфейсов и получить практический опыт их изучения.

Постановка задачи

Пользовательский интерфейс веб-сайта позволяет пользователю взаимодействовать с сайтом посредством графических образов.

Существуют разные подходы к разработке пользовательских интерфейсов, но в настоящее время господствующим является когнитивный подход или подход «извне-вовнутрь»: извне находится человек, дисплей и выводимая информация, а внутри – программа, алгоритм. Этот подход рассматривает пользователя как центральную фигуру взаимодействия с системой [1]. Опираясь на этот подход, корректно и эффективно разработанным интерфейсом будем считать тот интерфейс, который будет максимально удобен и понятен обычному пользователю. Чтобы сформулировать более точное определение, выделим два аспекта пользовательского интерфейса, которые непосредственно влияют и могут быть проанализированы программным образом: доступность веб-контента и использование и сочетания цветов на веб-странице.

Критерии доступности пользовательского интерфейса

Доступность веб-контента – это свойство веб-сайтов, приложений и других цифровых ресурсов быть доступными для всех пользователей, вне зависимости от их физических или когнитивных способностей [3].

В Российской Федерации требования по обеспечению доступности веб-контента для людей с ограниченными возможностями нормативно закреплены в ГОСТ Р 52872-2019, разработанном на основе актуального на тот момент международного стандарта WCAG или Руководства по обеспечению доступности веб-контента (англ. Web Content Accessibility Guidelines). В данной работе критерии были взяты из WCAG 2.2 от 5 октября 2023 года [2], последней актуальной на данный момент версии.

Для программной реализации были выбраны следующие критерии:

1. 1.1.1. Нетекстовой контент: весь представляемый пользователю контент сопровождается равноценной текстовой альтернативой.
2. 1.3.1. Информация и взаимосвязи: способ подачи информации, ее структура и взаимосвязи программным образом определяемы или доступны в текстовом виде.
3. 1.4.3. Контраст (минимальный): визуальное представление текста и изображений текста имеют контрастность не менее 4,5:1, для крупного текста контрастность не менее 3:1.
4. 1.4.6. Контраст (повышенный): визуальное представление текста и изображений текста имеют контрастность не менее 7:1, для крупного текста контрастность не менее 4,5:1.
5. 2.4.2. Заголовок страницы: веб-страницы снабжены заголовком, который описывает их содержание или назначение.
6. 3.1.1. Язык страницы: естественный язык по умолчанию каждой веб-страницы может быть определен программным образом.
7. 3.3.2. Ярлыки или инструкции: если контент требует ввода данных пользователем, предоставляются ярлыки или инструкции.

Критерии корректной работы с цветом в пользовательском интерфейсе

Цвет – мощный инструмент в руках веб-дизайнера. С помощью цвета можно подчеркнуть функционал веб-сайта и направить пользователя.

В отличие от доступности веб-контента, обеспечение которой контролируется стандартами и руководствами, четких правил по работе с цветом в пользовательских интерфейсах нет. Сформулированные критерии имеют рекомендательный характер.

Для программной реализации были выбраны следующие критерии:

1. Цвет кнопки подтверждения: цвет кнопки подтверждения не должен быть красным.
2. Цвет кнопки отмены: цвет кнопки отмены не должен быть зеленым.
3. Использование черного цвета: в дизайне элементов пользовательского интерфейса нежелательно использование чистого черного цвета.
4. Оформление одинаковых по функционалу элементов: интерактивные элементы пользовательского интерфейса, имеющие одинаковый функционал, должны иметь одинаковое оформление.
5. Оформление противоположных по функционалу элементов: интерактивные элементы пользовательского интерфейса, имеющие противоположный функционал, не должны быть одного цвета. [4]

Разработка программного модуля

Основными технологиями для разработки клиентской части веб-сайта являются HTML (язык гипертекстовой разметки документов), CSS (каскадные таблицы стилей) и JavaScript (язык программирования).

HTML используется для создания структуры веб-страниц и представляет собой набор тегов и атрибутов, которые объявляют и описывают каждый элемент на странице. CSS используется для описания внешнего вида веб-страницы, позволяет задавать стили для HTML-элементов, такие как цвет, размер шрифта, отступы и другое. JavaScript используется во frontend-части веб-сайта как язык сценариев, обеспечивая интерактивность веб-страницы и реализацию бизнес-логики.

Программный модуль определения ошибок был реализован на языке программирования JavaScript. Структура проекта программного модуля представлена на рисунке 1. Файлы в подпапке rules содержат в себе алгоритмы проверки правил. Основной файл программы – main.js. Файл utils.js содержит в себе вспомогательные функции. Файлы нужны index.css и index.html для веб-сайта, визуализирующего работу модуля.

Рис. 1. Структура проекта программного модуля

В модуле осуществляется парсинг проверяемых сайтов. При нажатии на кнопку «Загрузить сайт» и вводе URL-адреса, функция fetch() получает HTML-контент по указанному адресу, относительные ссылки заменяются на абсолютные и модифицированный контент вставляется в элемент iframe. Проверяемый сайт и связанные с ним ресурсы видны на веб-сайте, визуализирующем работу модуля.

Алгоритмы обработки правил разделены по категориям на отдельные js-файлы. Каждое правило – это отдельная функция, все правила хранятся в массиве rules в файле main.js. Проверка правил осуществляется за счет механизма проверки элементов DOM-дерева, полученного в результате парсинга сайта. DOM-дерево состоит из нод (node), которые представляют собой любой объект в дереве. Функция checkRules в файле main.js последовательно применяет все имеющиеся правила к DOM-дереву. Метод проверки правила выглядит следующим образом: с помощью селекторов выбираются нужные элементы (по HTML-тегам/атрибутам), получается массив нод, к каждой ноде или паре нод применяется проверка правила. Правила ссылаются на какой-либо атрибут или стили. Для проверки элементов на контрастность совершается рекурсивный обход DOM-дерева для поиска текстовых нод и извлечения у них родительского элемента, так как необходимо применить правило для текста и его фона.

В программный код модуля включены следующие интерактивные элементы: подсветка элементов с ошибкой, прокрутка на загруженном сайте к элементу с ошибкой, клик по заголовку критерия в правом блоке раскрывает подробную информацию об ошибке. Также в программном коде реализована проверка видимости элементов isVisibility() для исключения проверки правил этих элементов.

Результаты тестирования программного модуля

На рисунке 2 представлен скриншот основной страницы веб-сайта, иллюстрирующего работу модуля, в левом окне отображается анализируемый сайт, в правом окне отображается список найденных ошибок с подробной информацией о соответствующем критерии, сути ошибки, способе исправления ошибки и категории ошибки.

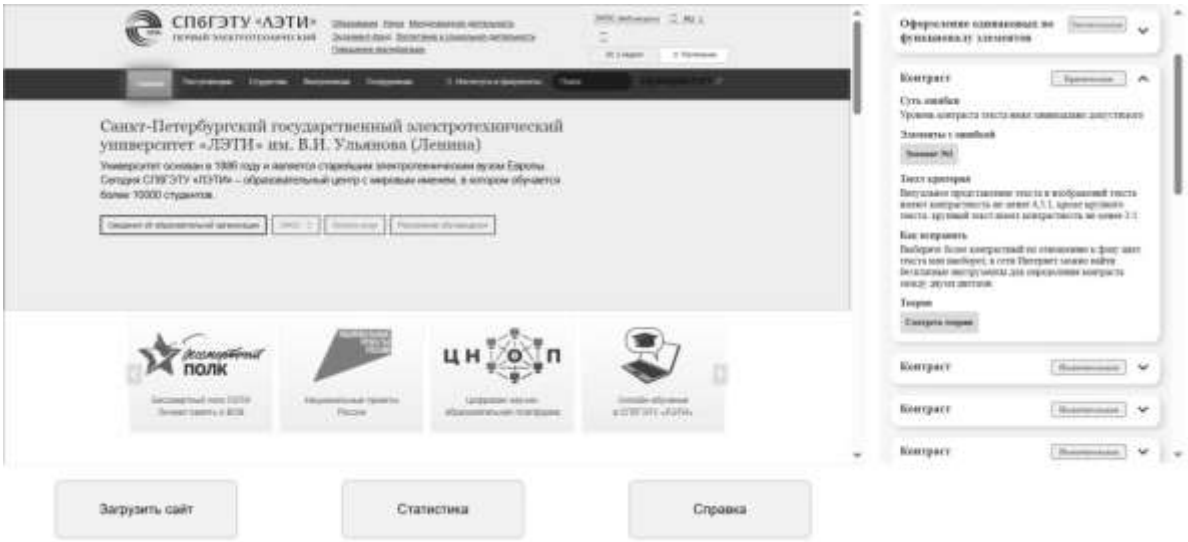


Рис. 2. Веб-сайт, демонстрирующий работу программного модуля

Работа программного модуля была протестирована на реальных веб-сайтах, а именно на главных страницах восьми высших учебных заведений Санкт-Петербурга. По результатам тестирования, на семи страницах были найдены ошибки по критерию «Нетекстовой контент» (отсутствовали текстовые описания изображений), на пяти страницах были найдены элементы со значением контраста ниже минимального значения, на трех страницах была найдена ошибка по критерию «Язык страницы» (отсутствовал программным образом определенный язык страницы) и на трех страницах была найдена ошибка по критерию «Оформление одинаковых по функционалу элементов».

Вывод

В результате выполнения работы был создан программный модуль, находящий ошибки на веб-страницах по заданным критериям корректности пользовательского интерфейса. Модуль показал свою работоспособность на реальных веб-страницах. Перспективами развития разработки будут: реализация анализа не только программного кода интерфейсов, но и их визуальной составляющей с помощью более мощных инструментов, например, нейронных сетей, а также добавление новых критериев для проверки по другим аспектам пользовательских интерфейсов, например, когнитивной нагрузке или сетке разметки элементов.

Список литературы

1. Лукин В.Н., Дзюбенко А.Л., Чечиков Ю.Б. Подходы к разработке пользовательского интерфейса // Программирование. 2020. №5. С. 16-24.
2. Руководство по обеспечению доступности веб-контента (WCAG) 2.2 [Электронный ресурс] URL: https://ifap.ru/ictdis/wcag22.htm /Дата обращения (06.04.2025)
3. Баланов А.Н. Прототипирование и разработка пользовательского интерфейса: оптимизация UX : учеб. пособие для вузов. Лань, 2024. 220 с.
4. Габриелян Т.О. Коммуникативный и мультимедийный дизайн. Графический пользовательский интерфейс : учеб. пособие. Крымский федеральный университет имени В. И. Вернадского, 2021. 166 с.

ПОДХОД К МОДЕЛИРОВАНИЮ ОТКАЗОУСТОЙЧИВЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

КУПРИЯНОВ Н. М.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В работе рассматривается подход к построению системы моделирования отказоустойчивых информационных систем.

Ключевые слова: Отказоустойчивость, информационная система, компьютерное моделирование

Введение

Создание безопасно функционирующих информационных систем (ИС) является одной из важнейших задач обеспечения безопасности отраслей экономики Российской Федерации. Действительно, в объектах критической инфраструктуры, которые содержат ИС разной сложности безусловно должны быть защищены, более того они должны проектироваться с учетом анализа уязвимостей в них. Отказоустойчивая система предназначена для гарантированного выполнения заданного набора функций (функциональной нагрузки) в течение заданного гарантированного промежутка времени. Это достигается за счет разных видов избыточности, в том числе структурной избыточности (в данном случае за счет использования нескольких параллельно работающих каналов). Данная проблематика рассматривается в работах отечественных и зарубежных авторов [1,6].

Целью данной работы является рассмотрение подхода к моделированию и построению отказоустойчивых структурно-избыточных информационных систем.

Функциональная нагрузка

Как уже говорилось выше, система предназначена для выполнения набора функций. Для каждой функции известно количество машинных тактов, необходимое для выполнения этой функции. Для удобства, время можно задавать и в единицах времени, тогда оно должно быть пересчитано в количество тактов, исходя из информации о производительности системы. Так же для каждой функции задается ее приоритет, а также время начала выполнения функции [2].

В случае, когда перед системой стоит задача максимального времени работы, то на этом множестве функций необходимо задать подмножество функций для обязательного выполнения.

Структура системы

Со структурной точки зрения система состоит из нескольких, параллельно работающих каналов (горячий резерв) и некоторого количества каналов холодного резерва. В горячем резерве канал выполняет все функции, и в нем могут происходить сбои. Холодный резерв отключен от работы и сбои в нем не происходят. Количество каналов горячего резерва может быть 3,5 или 7 (обязательно нечетное). Количество каналов холодного резерва не ограничено. В начале работы все каналы идентичны по своим характеристикам. Они имеют одинаковую производительность, в них с равной долей вероятности могут происходить сбои. В ходе работы системы характеристики каналов начинают меняться [2].

Для системы задается максимальный весовой коэффициент – МВК (например, не более 10). Этот коэффициент характеризует работоспособность. У каждого канала есть глобальный (ГВК) и текущий (ТВК) весовые коэффициенты. ТВК характеризует текущее состояние канала, а ГВК – состояние канала с учетом всей истории работы канала. В ходе работы каналов, в них могут происходить сбои. Частота возникновения сбоев задается либо в виде количества сбоев в канале в единицу времени, либо в виде экспоненциального закона с заданной интенсивностью сбоев. Частота сбоев одна и та же для всех каналов.

Режимы функционирования системы

Режим функционирования системы зависит от целей, поставленных перед ней и от порядка выполнения функций.

Перед системой может быть поставлена одна из следующих задач:

1. Выполнение всех функций. В этом случае система обязана выполнить все функции из функционального набора системы. После выполнения всех функций работа системы считается завершенной.
2. Максимальное время работы системы при заданном обязательном для выполнения множестве функций. В этом режиме система имеет право отказаться от выполнения каких-либо необязательных функций ради продолжения работы. В случае, если система не может выполнить какую-нибудь обязательную функцию, считается, что произошел отказ системы. Если система выполнила все функции, какие могла, она опять начинает выполнение функционального набора до полного отказа.

Порядок выполнения функций в системе может быть производным. В этом случае пользователь только указывает один из режимов выполнения функций:

1. Параллельное выполнение всех функций.
2. Последовательное выполнение функций. В этом случае функции могут быть отсортированы как по продолжительности выполнения, так и по приоритету. Так же существует режим работы, когда время начала выполнения каждой функции строго задано пользователем.

Описание функционирования системы

В ходе работы системы во всех каналах выполняются одинаковые функции. В системе возможно выполнение нескольких функций одновременно, при этом считается, что процессорное время делится поровну между всеми выполняемыми функциями. Время начала выполнения функций зависит от условий, рассмотренных выше. В случае, если система не выполняет функций предполагается, что каналы «спят» и сбоев в каналах не происходит.

Все время работы системы разбито на периодические контрольные точки. В этих точках происходит сравнение выходов каналов и определение сбоя в каком-либо канале. Определение выхода системы происходит по принципу большинства. Поэтому в системе

должно быть нечетное количество каналов. Если в системе в результате структурной деградации осталось только два канала, то работает режим «ведущий-ведомый». В данном случае, если произошло рассогласование, предпочтение отдается каналу, имеющему большую сумму ГВК и ТВК. Другой канал считается давшим сбой.

Обработка сбоев

Процедура обработки сбоев выполняет следующие действия:

- Оценивается время Δt между двумя последними сбоями в канале. Если это время меньше либо равно заданному времени T_1, то ТВК канала уменьшается на единицу. Заданием времени T_1 достигается игнорирование редких сбоев. Если в ходе работы системы ТВК какого-либо канала достигнет 0, то в этом случае уменьшается значение ГВК на единицу, а ТВК принимается равным МВК. Если в ходе работы системы время между сбоями Δt превысит заданное T_2 , то ГВК канала увеличивается на единицу (только в том случае, если он меньше МВК). Этим достигается восстановление доверия к каналу, в случае, если в нем долго не было сбоев.
- Если система работает на выполнение всех функций, то данный канал заменяется на канал из холодного резерва. Если же в холодном резерве нет работоспособных каналов, то система отключает данный канал. Если канал отключить нельзя, то считается, что система дала отказ.
- Если система работает на максимальное время работы, и при этом она может выполнить какую-либо функцию из обязательного набора функций, то заданный канал отключается. Если же система не может выполнить ни одной функции, то заданный канал заменяется на канал из холодного резерва. Если в холодном резерве нет каналов, то система оказала.

Описание системы моделирования

Система моделирования должна позволять пользователю вводить все исходные данные, следить за ходом эксперимента и просматривать в удобной форме полученные результаты работы и статистику [5,6].

Входные данные

Входные данные включают в себя:

- Описание системы:
  - Количество каналов в «горячем резерве» (3,5,7);
  - Количество каналов в холодном резерве (больше либо равно 0);
  - Максимальный весовой коэффициент;
  - Время выполнения одного машинного такта (частота процессора);
  - Время между контрольными точками;
  - Время на обработку контрольной точки при нормальной работе;
  - Время обработки отказа;
  - Время ввода канала из холодного резерва;
  - Время T_1, задающее игнорируемый период сбоев;
  - Время T_2, задающее время восстановления доверия к каналу;
- Режим работы системы по типу задания:
  - Выполнение всех функций;
  - Максимальное время работы при заданном наборе функций;
- Режим работы системы по началу выполнения функции:

- Фиксированное начало выполнения функции;
- Одновременное начало выполнения всех функций;
- Последовательное начало выполнения функций:
  - Тип сортировки: по приоритету либо по продолжительности;
  - Направление сортировки;
- Задание частоты сбоев в виде:
  - Количества сбоев в канале в единицу времени;
  - Интенсивность сбоев при экспоненциальном законе;
- Описание функционального набора. Для каждой функции:
  - Время выполнения;
  - Приоритет;
  - Время начала выполнения (если необходимо);
  - Флаг необходимости выполнения (если необходим);

Результаты работы

Система моделирования должна динамически отображать состояние системы. Пользователь должен в любой момент иметь возможность получить доступ к любой переменной системы для чтения. Должна выводится следующая информация:

- Состояние каналов:
  - ГВК;
  - ТВК;
  - Количество сбоев в канале;
  - Время, прошедшее после последнего сбоя;
  - Канал в холодном, или горячем резерве;
- История событий (какие события и когда произошли)
- Функциональный набор:
  - Статус функции (выполненная, невыполняемая функция, функция выполняется в данный момент)

После окончания моделирования получаем следующие данные:

- Время работы системы;
- Количество циклов, выполненных системой;
- Количество сбоев в системе;
- Количество выполненных функций. При работе на максимальное время работы – сколько раз каждая функция была выполнена.

Заключение

На основании анализа и полученных результатов можно сделать следующие выводы:

5. Предложен алгоритм реконфигурации структурно избыточной ИС с учетом катастрофических и перемежающихся отказов.
6. Определен подход к созданию системы моделирования отказоустойчивых ИС.

Список литературы

1. И. Б. Шубинский, Е. Н. Розенберг «Функциональная безопасность систем управления на железнодорожном транспорте», 2023, ISBN: 978-5-9729-1553-8, 361 с.

2. Israel Koren, C. Mani Krishna Fault tolerant systems. _ Elsevier, 2007, ISBN 0-12-088525-5/

3. Baha Al-Shaikh, Simon G. Stacey, Essentials of Equipment in Anaesthesia, Critical Care, and Peri-Operative Medicine (2017), p. 247.

4. Dubrova, E. (2013). "Fault-Tolerant Design", Springer, 2013, ISBN 978-1-4614-2112-2
5. Reliability evaluation of some fault-tolerant computer architectures. Springer-Verlag. November 1980. ISBN 978-3-540-10274-8.
6. Herzberg, Amir; Shulman, Haya (2012). "Oblivious and Fair Server-Aided Two-Party Computation". 2012 Seventh International Conference on Availability, Reliability and Security. IEEE. pp. 75–84. doi:10.1109/ares.2012.28. ISBN 978-1-4673-2244-7. S2CID 6579295

РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ УЧЁТА ПОТРЕБЛЕНИЯ КАЛОРИЙ С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

МЕРЗЛЯКОВ И.А., СИНЕВ В.Е.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В. И. Ульянова (Ленина)

Аннотация. В статье представлено исследование, посвященное разработке и оценке эффективности веб-приложения для учета калорий с использованием искусственного интеллекта DeepSeek. Основной целью работы является создание инструмента, который автоматизирует процесс учета питания, предоставляет точные расчеты калорийности и персонализированные рекомендации по питанию.

Ключевые слова: учет калорий, искусственный интеллект, DeepSeek, персонализированные рекомендации, здоровое питание

Введение

Здоровый образ жизни и правильное питание становятся неотъемлемой частью повседневной жизни, что увеличивает спрос на инструменты для контроля рациона. Ручной подсчет калорий часто оказывается трудоемким и неточным, что подчеркивает необходимость автоматизированных решений. Современные технологии искусственного интеллекта (ИИ), такие как DeepSeek, позволяют создавать интеллектуальные системы для анализа данных, распознавания продуктов и прогнозирования их калорийности.

Цель работы — разработка и оценка эффективности веб-приложения, интегрирующего ИИ для автоматизации учета калорий и предоставления персонализированных рекомендаций.

Обзор существующих решений

Популярные приложения для учета калорий, такие как MyFitnessPal и FatSecret, предлагают функции ручного ввода данных, базы продуктов и интеграцию с фитнес-трекерами. Однако они имеют ряд ограничений:

1. Неточность данных в базах продуктов.
2. Ограниченная персонализация рекомендаций.
3. Сложность ввода данных.

DeepSeek, благодаря высокой точности обработки естественного языка и способности анализировать индивидуальные предпочтения, предлагает более совершенный подход.

Методология

Архитектура приложения включает следующие компоненты:

- Frontend: Vue.js для пользовательского интерфейса.
- Backend: Node.js (Express.js) и PostgreSQL для хранения данных.

- Микросервисы:
  - Food Service для работы с базой продуктов.
  - Auth Service для аутентификации.
  - User Service для управления данными пользователей.
- DeepSeek API: Для анализа текстовых описаний продуктов и генерации рекомендаций.

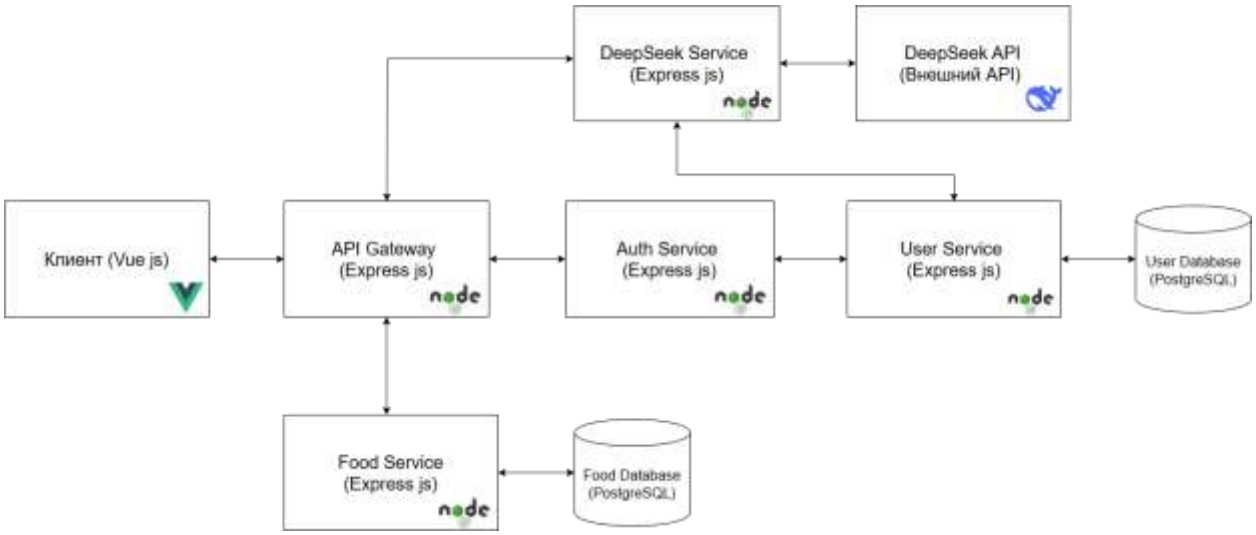


Рис. 1. Схема архитектуры приложения

Функциональные модули:

1. Расчет дневной нормы калорий на основе данных пользователя.

Рис. 2. Блок информации о пользователе с рассчитанной дневной нормой калорий

2. Добавление продуктов в рацион через “умный” текстовый ввод.

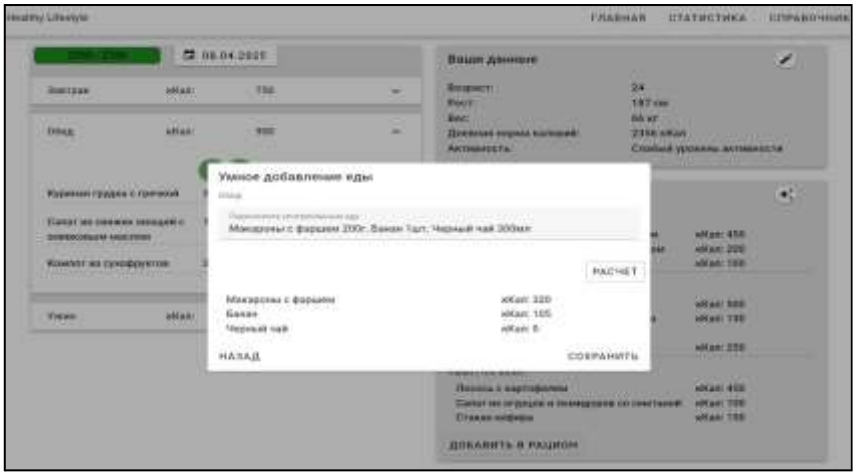


Рис. 3. Форма для расчета калорийности введенных вручную продуктов на основе DeepSeek API

3. Генерация персонализированного плана питания.

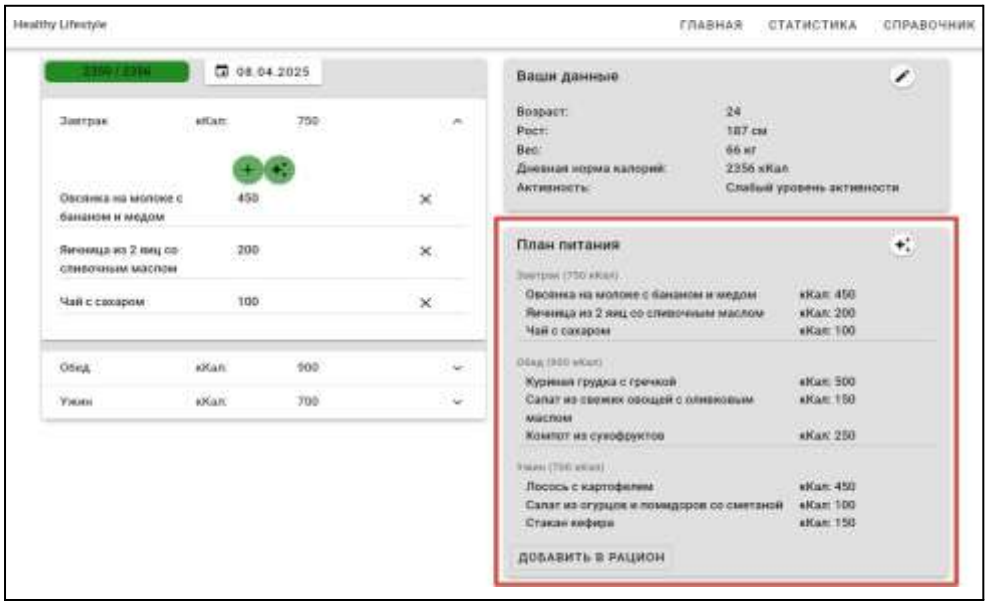


Рис. 4. Модуль генерации плана питания

4. Визуализация статистики потребления калорий.

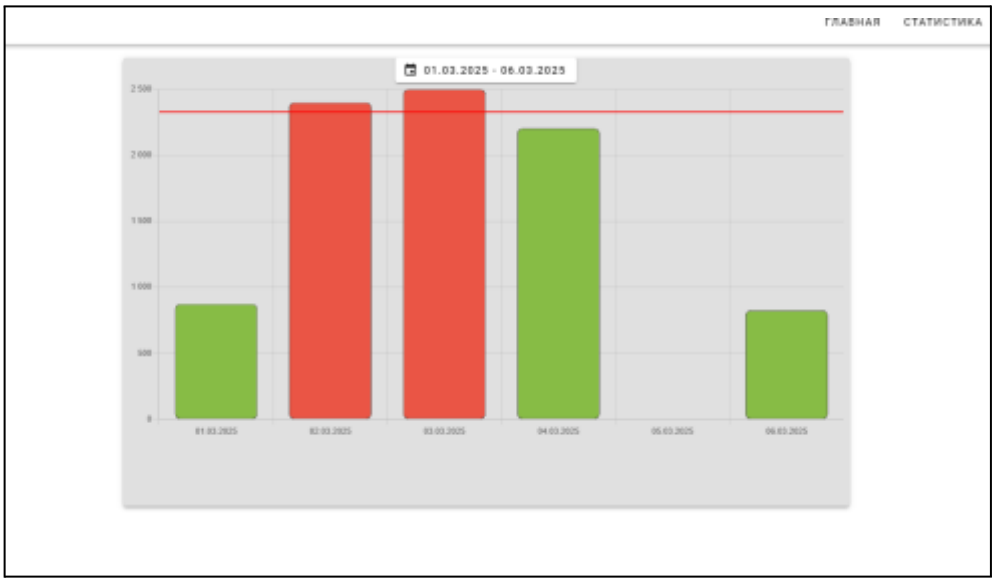


Рис. 5. Окно статистики потребления калорий

Анализ данных

Точность расчетов калорийности продуктов сравнивалась с данными авторитетных источников (health-diet.ru, beregifiguru.ru).

Результаты:

- Калорийность: Среднее отклонение — 2.64% (health-diet.ru) и 5.38% (beregifiguru.ru).
- Белки: Среднее отклонение — 6.71% и 7.75%.
- Жиры и углеводы: Наибольшие отклонения (до 15.71% и 23.17%) связаны с различиями в рецептах и методах расчета.

Заключение

Разработанное приложение демонстрирует высокую точность и удобство использования, превосходя существующие аналоги по персонализации рекомендаций.

Перспективы развития включают интеграцию учета физической активности с использование DeepSeek API и создание Telegram-бота для взаимодействия с пользователями.

Приложение представляет собой эффективный инструмент для поддержания здорового образа жизни с использованием современных технологий ИИ.

Список литературы

1. Приложение для учета калорий и контроля питания [Электронный ресурс] - Режим доступа: https://www.myfitnesspal.com/ (дата обращения 21.02.2025)
2. Приложение для учета калорий и планирования диет [Электронный ресурс] - Режим доступа: https://www.fatsecret.com/ (дата обращения 25.02.2025)
3. Российский ресурс с информацией о калорийности продуктов [Электронный ресурс] - Режим доступа: https://health-diet.ru/ (дата обращения 28.02.2025)
4. Сайт с данными о калорийности и питательной ценности продуктов [Электронный ресурс] - Режим доступа: https://beregifiguru.ru/ (дата обращения 28.02.2025)
5. Официальная документация по фреймворку Vue.js для разработки пользовательских интерфейсов

[Электронный ресурс] - Режим доступа: https://vuejs.org/ (дата обращения 28.02.2025)
6. Официальная документация по СУБД PostgreSQL [Электронный ресурс] - Режим доступа: https://www.postgresql.org/docs/ (дата обращения 01.03.2025)

АНАЛИЗ ЭЛЕМЕНТОВ НАВИГАЦИИ СТАНЦИИ МЕТРО "МОСКОВСКАЯ" ПЕТЕРБУРГСКОГО МЕТРОПОЛИТЕНА

К.О. Моисеев, А.О. Трофимова

[1]*Санкт-Петербургский государственный электротехнический университет«ЛЭТИ» им. В. И. Ульянова (Ленина)*

Аннотация. Статья посвящена анализу и совершенствованию навигационной системы на станции метро «Московская» в Санкт-Петербурге Проведён обзор существующих типов навигационных схем и выявлены их недостатки, включая избыточность информации, дублирование, визуальные помехи и недостаточную читаемость. В результате работы были разработаны новые схемы, представлены рекомендации Представленные рекомендации направлены на улучшение пользовательского опыта и повышение эффективности навигации для пассажиров, особенно в условиях интенсивного пассажиропотока.

Ключевые слова: схема, информационный стенд, навигация, метрополитен Санкт-Петербурга.

«Московская» — станция закрытого типа («горизонтальный лифт») глубокого заложения (глубина ≈ 29 м). По краям платформ установлены 52 платформенные раздвижные двери. Выходы со станции располагаются в обоих торцах, в каждом выходе по три эскалатора. Основная сложность в ориентировании на этой станции – наличие двух вестибюлей и 8 выходов.

Цель работы - улучшить навигацию на станции метро Московская и упростить нахождение нужного выхода на улицу.

Задачи:

1. Рассмотреть текущую систему навигации на станции метро Московская.
2. Выявить слабые места в текущей системе навигации.
3. Выявить недостатки в информационных стендах и схемах.
4. Выработать комплекс мер по улучшению навигации на станции.

В среднем, пассажиропоток на станции метро Московская составляет 76 тысяч человек в день. В первую очередь для пассажиров станция метро Московская – пересадочный пункт на пути между городом и аэропортом Пулково. По статистике за тот же период автобусами маршрута 39 (между городом и аэропортом) в среднем пользовались 23 тысячи пассажиров.

В том же направлении ходят автобусы в сторону парка Героев, торговых центров, Университета гражданской авиации и пригорода. Основная остановка перечисленных маршрутов находится у выхода 1 вестибюля 2. Входом 4 вестибюля 2 преимущественно пользуются сотрудники 156 компаний, размещенных в доме Советов, а также туристы, приезжающие посмотреть на него и фонтаны в летнем сезоне. Остальные выходы менее востребованы, но необходимы, поскольку также являются единственной возможностью оказываться по разные стороны широкого Московского проспекта.

Недостатки навигации на станции

Основная задача навигации на станции метро Московская – направить человека к автобусам в аэропорт «Пулково», однако заметны следующие недостатки:

- 2 огромных ремонтных короба посередине платформы закрывают собой вид на схемы и затрудняют поиск нужного пути;

- лишняя информация на навигационных схемах отвлекают от основной задачи – выйти из метро в нужном месте;
- отсутствие единообразия в схемах;
- много дублирующийся информации на разных типах схем;
- обилие яркой и отвлекающей рекламы в непосредственной близости с навигационными схемами.

На станции метро можно выделить несколько типов навигационных схем, однако сосредоточимся на тех, которые помогают покинуть станцию и оказаться в нужном человеку месте.

Тип 1: Схема с информацией об остановках общественного транспорта рядом с каждым из выходов из метро.

Недостатки:

- отсутствие наглядности (выходы представлены вертикальными списками);
- сложность поиска интересующего маршрута (3 из 98 представленных на схеме номеров маршрутов);
- дублирующаяся информация (60% маршрутов указаны 2-3 раза);
- много пустого места на схеме.

Тип 2: Новая схема с информацией о том, как пройти к тому или иному выходу из метро. Данная схема появилась на станции недавно и призвана решить накопившиеся проблемы с поиском выхода из метро, она более наглядная и конкретная относительно предыдущей, однако представлена на стенах реже и меньше привлекает внимание пассажиров.

Недостатки:

- на схеме слишком большую часть полезного пространства занимает текст, это мешает быстрому пониманию схемы и рассеивает внимание.

Тип 3: Главный информационный стенд. На нем размещена карта местности над станцией, информация о наземном транспорте и правила пользования метро.

Тип 4: Онлайн-табло аэропорта «Пулково». Размещается на цифровом экране в непосредственной близости к выходам из метро и содержит информацию о прилете\вылете самолетов из аэропорта «Пулково».

Недостатки:

- фотография автобуса экспресса в аэропорт «Пулково» занимает четверть пространства на экране;
- неправильно расставлены приоритеты информации;
- кривая верстка онлайн-табло;
- неиспользуемый столбец «Примечание»;
- отправленные рейсы занимают пространство будущих (имеющих значение для пассажира, следующего в аэропорт).

Тип 5: Схемы линий метро и синей ветки, в частности. Их на станции очень много (32 из 64 схем). Они стандартны для метро в целом, поэтому их улучшение выходит за рамки этой статьи [2].

Ограничения

В правилах пользования Петербургским метрополитеном с изменениями и дополнениями от 31 мая 2024 года [1] регламентированы требования к информированию пассажиров, находящихся в вагонах, на станциях и межстанционных переходах. Перечислим требования, относящиеся к типам и наполнению печатной навигации на станциях:

1. Выходы со станций должны быть пронумерованы.
2. Перевозчик обязан обеспечить размещение на станциях: указателей с наименованием станции; схем линий метрополитена; правил пользования метрополитеном; указателей направлений и номеров выходов со станции; информации о времени работы станции;
3. Количество предусмотренных к размещению на станциях схем (карт), иных информационных материалов, а также места их размещения должны определяться перевозчиком.
4. Информирование, осуществляемое в соответствии с пунктом 2 должно дублироваться метрополитеном на английском языке.

Разработка новых навигационных схем

Внимание было уделено схемам 1, 2 и 4 типа. Схемы 1-го и 2-го типа были объединены в одну.

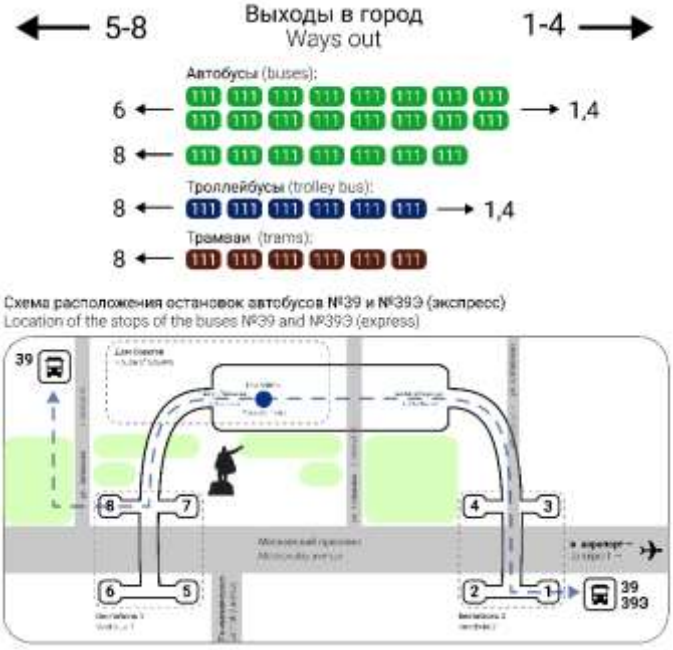


Рис. 2. Пример верстки новой схемы выходов из метро

На данной схеме отражены основные топографические элементы местности, направления движения в соответствии с остальными указателями станции, вестибюли и все их выходы, номера маршрутов общественного транспорта с ближайшими к остановкам выходами, и основной акцент сделан на направлении к автобусам в аэропорт.

Экран с онлайн-табло аэропорта Пулково (схема 4 типа) также было доработано, чтобы улучшить основную его функцию – размещение информации о расписании самолетов:

- представлена только необходимая для пассажира;
- добавлена цветовая индикация для упрощения определения состояния рейса;
- исправлена верстка.

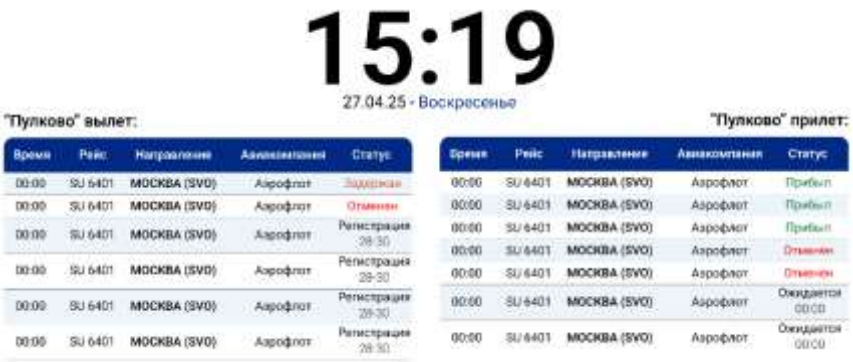


Рис. 3. Пример верстки нового экрана онлайн-табло

Рекомендации по общему улучшению навигации

1. Сократить количество информационных схем на платформе (с 64 до 32).

Когда информационных схем много, и они многократно повторяют одно и то же – человек начинает теряться.

2. Убрать яркие рекламные экраны.

Яркие экраны отвлекают от важной информации и портят внешний вид станций, особенно учитывая, что для многих гостей города Московская – первая станция метро, которую они встречают в своем путешествии.

3. Более явно выделять информационные стенды.

Только один тип информационного стенда отмечается знаком «I», выступающим от стены. Остальные не выделяются, мало отличаются друг от друга и теряются на фоне яркой и большей по размеру рекламы.

4. Избавление от рабочих коробов на станции.

Во-первых, коробы портят внешний вид станции, о важности данной проблемы было упомянуто ранее. Во-вторых, они сдавливают поток пассажиров к краям станции. Это затрудняет остановку для взаимодействия со схемами. В-третьих, они перекрывают обзор на все информационные указатели. Если отсутствует возможность убрать коробы совсем, на них нужно крупно указать направления движения.

5. Более крупное и яркое обозначение выходов из метро в подземном переходе.

На поверхности, уже после выхода из вестибюля, человек уже должен понимать, к какому выходу он направляется. Сами таблички с номерами располагаются почти у лестниц выходов, заметить их можно только когда человек уже прошел мимо 70% ширины прохода на выход.

6. Навигация внутри вагона метро.

Вагоны метро «привязаны» к линиям, по которым они передвигаются. Поэтому в них можно заранее разместить информацию о направлениях движения на станциях метро соответствующей линии, чтобы, выходя из вагона, пассажир уже знал, в какую сторону ему идти. Аналогичная навигация уже используется в Московском метрополитене.

Примеры реализации всех решений представлены на презентации в рамках выступления на конференции.

Заключение

Проведённый анализ системы навигации на станции метро «Московская» выявил недостатки, затрудняющие ориентирование пассажиров, особенно тех, кто спешит к автобусам в аэропорт «Пулково». Были выявлены и устранены избыточные и дублирующие элементы в схемах, переработаны основные навигационные форматы, включая объединение схем 1-го и 2-го типов и модернизацию онлайн-табло аэропорта. Внедрение

новых схем сделало информацию более лаконичной и удобной для быстрого восприятия. Предложенные меры по улучшению визуального восприятия навигации, сокращению рекламной нагрузки и устранению физических препятствий направлены на повышение комфорта и безопасности пассажиров. Эти шаги особенно важны, учитывая высокий пассажиропоток и значимость станции как ключевого транспортного узла между городом и аэропортом.

Список литературы

1. Постановление Правительства Санкт-Петербурга от 29 сентября 2020 г. № 776 «Об утверждении Правил пользования Петербургским метрополитеном» (с изменениями и дополнениями от 25 ноября 2021 г., 31 мая 2024 г.) // Собрание законодательства Санкт-Петербурга. 2020.
2. Моисеев К. О., Трофимова А. О. Проектирование схемы линий Петербургского метрополитена // Наука настоящего и будущего: XII науч.-практ. конф. с междунар. участием для студентов, аспирантов и молодых ученых: сб. матер. конф. В 2 т. Т. 2. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2024. С. 340-344.

МЕТОДИКИ МОДЕЛИРОВАНИЯ УГРОЗ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В СИСТЕМАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

НАЛИМОВ К.Г.

Санкт-Петербургский политехнический университет Петра Великого

Аннотация. В работе исследуются методы и техники моделирования угроз информационной безопасности в информационных системах с использованием технологий и алгоритмов искусственного интеллекта. Рассматриваются как классические методы моделирования угроз, основанные на статическом моделировании и поведенческом анализе пользователей информационных систем, так и методы, базирующиеся на нейросетевых технологиях и алгоритмах глубокого машинного обучения. Ключевой акцент делается на обнаружение угроз и выявление аномалий в информационных системах искусственного интеллекта.

Ключевые слова: искусственный интеллект, информационная безопасность, противодействие угрозам, методы моделирования угроз информационной безопасности

На сегодняшний день область разработки систем и цифровых продуктов искусственного интеллекта является наиболее перспективным и востребованным направлением индустрии информационных технологий. Стремительный рост технологий искусственного интеллекта влечет за собой появление принципиально новых угроз информационной безопасности и как следствие необходимость пересмотра методик моделирования угроз информационной безопасности характерных для систем искусственного интеллекта. Данными фактами обусловлена актуальность данного исследования. На рис. 1 представлены показатели роста информационных систем и продуктов искусственного интеллекта [1].

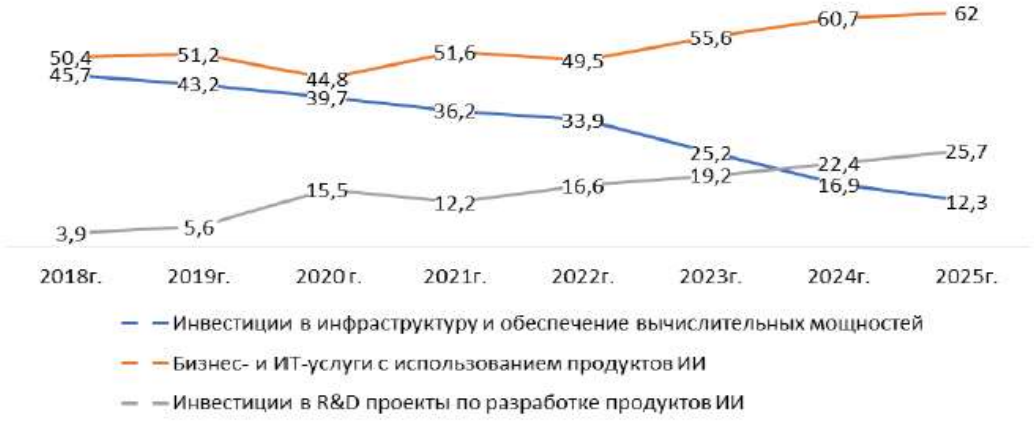


Рис. 1. Показатели роста систем искусственного интеллекта в РФ

Цель исследования заключается в обзорном анализе систем искусственного интеллекта и выборе оптимальной методики моделирования угроз информационной безопасности.

Для моделирования угроз информационной безопасности системы искусственного интеллекта необходимо классифицировать объект защиты по степени влияния алгоритмов искусственного интеллекта на приятие решений в автоматизированной информационной системе. Классификация информационных систем по степени влияния алгоритмов искусственного интеллекта на процессы принятия решений приведена в таблице 1.

Таблица 1

Классификация информационных систем с применением технологий искусственного интеллекта

№	Модель влияния ИИ	Описание
1	Вспомогательная модель	Алгоритмы искусственного интеллекта используются в качестве вспомогательного инструментария. Принятие решение в работе процессов информационной системе остается за естественным интеллектом, за человеком.
2	Гибридная модель	Искусственный и естественный интеллект выступают в качестве партнеров. Алгоритмы искусственного интеллекта выполняют сложные задачи и обрабатывают большие массивы данных в информационной системе, но окончательное принятие решений остается за человеком.
3	Инструментальная модель	Алгоритмы искусственного интеллекта реализуют полный цикл процессов в информационной системе, естественный интеллект выступает в качестве контролера реализуемых ИИ процессов.
4	Автономная модель	Искусственный интеллект выступает в качестве автономного исполнителя всех процессов в информационной системе. Участие человека в работе информационной системы и принятии решений не требуется.

Для выявления и предотвращения угроз информационной безопасности и кибербезопасности для информационных систем искусственного интеллекта с различными моделями влияния на принятие решений могут быть применены следующие методы моделирования угроз, представленные в таблице 2 [2].

Таблица 2

Методы моделирования угроз информационной безопасности

№	Метод	Описание
1	Метод с применением искусственных нейронных сетей	Моделирование с использованием нейронных сетей, использующие большие массивы данных об угрозах и аномалиях информационных систем в качестве обучающего набора данных, которые позволяют устанавливать сложные зависимости для прогнозирования угроз.
2	Метод генетических алгоритмов и алгоритмов глубокого машинного обучения	Моделирования с применением эволюционных принципов и алгоритмов для выявления аномалий в информационных системах и выбора оптимальных сценариев реагирования.
3	Методы моделирования на основе поведенческого анализа	Метод основан на использовании классических инструментов поведенческого анализа, анализа сетевого трафика и общего состояния информационной системы для выявления отклонений от эталонных значений и заданных политик информационной безопасности.
4	Методы статистического моделирования	Метод моделирования с использованием статистических инструментов и моделей для прогнозирования и выявления аномалий и угроз информационной безопасности в информационной системе.

Методы моделирования угроз информационной безопасности с применением искусственных нейронных сетей, генетических алгоритмов и алгоритмов глубокого машинного обучения являются эффективными и перспективными методами для прогнозирования угроз, уязвимостей и аномалий в информационных системах, но требуют качественного экспертного анализа результатов моделирования и проверки достоверности полученных результатов. Следует отметить, что нейросетевая модель прогнозирования представляет собой адаптивную модель интеллектуального анализа данных и принятия решений. Адаптивность модели заключается в возможности автоматической корректировки параметров нейронной сети, что позволяет модели обновляться и аппроксимировать новые закономерности в поведении временного ряда [3].

Однако применения искусственных нейронных сетей для обеспечение информационной безопасности в моделях искусственного интеллекта имеет критическое значение, сопоставимое с важностью кибербезопасности. Модели искусственного интеллекта, случайно раскрывающие личную информацию в процессе анализа угроз, могут сделать системы уязвимыми для будущих неожиданных атак. Ключевая задача заключается в том, чтобы обучить модели искусственного интеллекта эффективно справляться с различными кибератаками, при этом оптимизируя время реагирования и сохранение конфиденциальности [4].

Каждый тип угроз информационный безопасности для системы искусственного интеллекта является уникальным и может варьироваться в зависимости от целей и назначения конкретной системы [5]. В данной работе были выделены три ключевые угрозы для системы искусственного интеллекта:

- искажение логики работы модели искусственного интеллекта;
- извлечение модели искусственного интеллекта (угроза нарушения целостности и доступности модели);

- нарушение конфиденциальности модели искусственного интеллекта.

Наиболее целесообразным является комбинированный подход к использованию всех выше представленных методов моделирования, учитывающий характер применения технологий искусственного интеллекта и степень влияния алгоритмов искусственного интеллекта на принятие решений в той или иной информационной системе.

Список литературы

1. Пантелеева Т.А. Возможности и угрозы использования искусственного интеллекта в бизнес-форсайте российских компаний // Проблемы рыночной экономики. 2021. № 1. С. 131-148.
2. Щербаков А.Е. Исследование применения искусственного интеллекта и машинного обучения в области кибербезопасности: техники обнаружения аномалий и предотвращения угроз // Международный научный журнал «Вестник науки». 2023. Т1. №7(64). С. 151–155.
3. Катасёв А. С., Катасёва Д. В., Кирпичников А. П. Нейросетевое прогнозирование инцидентов информационной безопасности предприятия // Вестник технологического университета. 2013. Т18. №3. С. 215-218.
4. Беляева Т.А., Микрюков А.А. Нейросетевое прогнозирование инцидентов информационной безопасности // Международный студенческий научный вестник. 2023. № 6. URL: https://eduherald.ru/ru/article/view?id=21424 (дата обращения: 03.05.2025).
5. Путято М. М., Макарян А.С. Кибербезопасность как неотъемлемый атрибут многоуровневого защищенного киберпространства // Прикаспийский журнал: управление и высокие технологии. 2020. № 3 (51). С. 94–102.

МЕТОДЫ ПОВЫШЕНИЯ КАЧЕСТВА ОПИСАНИЙ УЯЗВИМОСТЕЙ ПО НА ОСНОВЕ ДАННЫХ ИЗ ОТКРЫТЫХ ИСТОЧНИКОВ

СЕМЕЙКИН С.А., КУЛАГИН М.В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Аннотация. Статья предлагает подробный анализ проблемы неполноты описаний уязвимостей в отечественном программном обеспечении. В статье рассматриваются причины недостаточной информативности записей об уязвимостях и предлагаются методы их улучшения за счёт сопоставления с иностранными источниками. Особое внимание уделяется таким инструментам, как прямое связывание с CVE, применение технологий обработки естественного языка и сравнение CVSS-векторов. В статье рассмотрен пример реальной уязвимости ПО, которая может быть дополнена с использованием описанных методов. Статья будет полезна специалистам по информационной безопасности и разработчикам, заинтересованным в повышении качества отечественных баз уязвимостей.

Ключевые слова: уязвимости, отечественное программное обеспечение, CVE, информационная безопасность, сопоставление уязвимостей, CVSS, описание уязвимостей, машинное обучение, обработка текста, база данных уязвимостей, повышение качества.

Введение

С активным развитием импортозамещённого ПО (Astra Linux, RedOS, Jatoba и др.) в России остро встаёт вопрос качества описаний уязвимостей. В отличие от иностранного ПО, отечественные продукты зачастую сопровождаются неполной информацией: нет детальных описаний, оценок уязвимостей (метрик CVSS, EPSS, Cisa), методов эксплуатации (эксплойтов) и даже прямых ссылок на известные системы идентификации (CVE, OSV, GHSA) [1-5].

Но поскольку многие российские продукты основаны на открытых зарубежных решениях, уязвимости в них, как правило, имеют прямые аналоги в иностранных системах.

Это позволяет легко дополнять и обновлять данные об уязвимостях, сопоставляя российские продукты с их зарубежными аналогами.

В контексте статьи под **описаниями уязвимостей** понимаются структурированные записи, содержащие сведения о конкретной уязвимости: её идентификатор (например, CVE-1999-0001), краткое описание, потенциальные последствия, уровень критичности (CVSS), вероятность эксплуатируемости (EPSS), ссылки на исходные коды, обновления и другие источники. Эти описания служат основой для понимания природы уязвимости и оценки её влияния на систему.

Уязвимости, обнаруженные в зарубежных проектах, часто напрямую применимы к российским продуктам, основанным на них:

- Astra Linux основана на Debian, следовательно многие уязвимости из Debian релевантны и для Astra
- RedOS — производная ОС от Red Hat Enterprise Linux
- СУБД Jatoba основана на PostgreSQL
- Браузеры, например, Яндекс.Браузер, используют ядро Chromium

Во многих международных бюллетенях и CVE присутствуют ссылки на родственные проекты, обновления, обсуждения и даже git-коммиты. При этом российские вендоры часто не указывают эти связи, из-за чего:

- Уязвимости дублируются
- Описания уязвимостей не содержат метрик (CVSS, EPSS, Cisa)
- Описания уязвимостей не содержат рекомендаций по их устранению

Способы улучшения

Рассмотрим способы, которые могут улучшить качество и количество определяемых уязвимостей, путем агрегации нескольких источников.

- **Использование данных производителей ПО**

Из известных отечественных продуктов своими базами данных уязвимостей могут похвастаться лишь RedOs и AstraLinux, остальные вендоры предоставляют минимальную информацию о связи версий используемых ядер или зависимостей. Таким образом, ориентируясь на полученную информацию и используя существующие уязвимости, можно формировать дополнения к ним и повышать эффективность анализаторов, путем сопоставления версий ПО.

- **Прямое связывание записей**

Частым случаем являются уже проставленные ссылки в описаниях уязвимостей, так, например, отечественная база данных уязвимостей ФСТЭК [6] имеет ссылки на идентификаторы CVE. Используя оба источника можно значительно расширить описание, добавив в него информацию об известных эксплоитах, исправлениях, а также метриках.

- **Анализ текста описания уязвимостей**

С помощью современных методов NLP и применения моделей машинного обучения можно находить схожие описания, которые до этого не были связаны с друг другом, таким образом расширяя отечественные базы данных.

- **Сравнение метрик**

Зачастую, у связанных уязвимостей прослеживаются одинаковые метрики, ведь они строятся по одним и тем же правилам, и из-за изменения надстройки над ядром

программного продукта не меняется суть уязвимости и полученная метрика остается такой же.

Используя CVSS вектор третьей или четвертой версии и сравнивая его вектором похожей уязвимости можно выявлять в них зависимости и приравнивать такие уязвимости при большой доле соответствия.

Практическое применение

Рассмотрим описание уязвимости с идентификатором BDU:2022-04971. Исходя из информации БДУ ФСТЭК, PostgreSQL версий до 15 Beta 3 является уязвимой.

В таблице 1 указано сопоставление версий Jatoba и PostgreSQL [7].

Таблица 1

Жизненный цикл версий СУБД Jatoba

Версия Jatoba	*Ядро PostgreSQL*	*Срок окончания поддержки*
1	*11*	*18.11.2023*
4	*14*	*12.11.2026*
5	*15*	*11.11.2027*
6	*16*	*09.11.2028*

Ориентируясь на таблицу жизненного цикла версий СУБД Jatoba для ядра PostgreSQL 15 нужно рассматривать версию Jatoba 5. Таким образом можно сделать вывод о том, что BDU:2022-04971 является применимой по отношению к русской СУБД Jatoba, несмотря на то, что этой информации в опубликованной уязвимости не было.

Вывод

Улучшение описаний уязвимостей отечественного программного обеспечения возможно за счёт активного использования информации из международных источников. Поскольку многие российские продукты основаны на зарубежных решениях, уязвимости в них часто имеют прямые аналоги, что позволяет расширить и уточнить их описание. Прямое связывание с CVE, применение методов обработки естественного языка и анализ CVSS-векторов позволяют устранить дублирование, добавить недостающие метрики и уточнить пути устранения уязвимостей. Такой подход существенно повышает качество и практическую ценность отечественных баз данных уязвимостей.

Список литературы

1. Описание системы оценки EPSS [Электронный ресурс] - URL: https://www.first.org/epss/user-guide
2. Оценка уязвимостей CVSS [Электронный ресурс] - URL: https://habr.com/ru/companies/pt/articles/266485/
3. Официальная странница БДУ NVD [Электронный ресурс] - URL: https://nvd.nist.gov/
4. Официальная странница БДУ OSV [Электронный ресурс] - URL: https://osv.dev/
5. Официальная странница БДУ Github [Электронный ресурс] - URL: https://github.com/advisories
6. БДУ ФСТЭК [Электронный ресурс] - URL: https://bdu.fstec.ru/vul
7. Описание СУБД Jatoba [Электронный ресурс] - URL: https://jatoba.ru/

СТЕГАНОГРАФИЯ В НЕСЖАТОМ ВИДЕОПОТОКЕ С ИСПОЛЬЗОВАНИЕМ ДИСКРЕТНОГО КОСИНУСНОГО ПРЕОБРАЗОВАНИЯ

Солонухина А.Л., Шкляр Е.В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И.Ульянова (Ленина)

Аннотация. В статье рассмотрена стеганографическая система, использующая в качестве контейнера несжатый видеопоток и основанная на методе дискретного косинусного преобразования. Определены компоненты рассматриваемой стеганографической системы и приведено описание используемого метода. На основе анализа литературы проведено сравнение используемого метода с другими классическими методами видеостеганографии по трем показателям: скрывающая способность, незаметность и устойчивость, – а также выявлены преимущества и вызовы использования этого метода в рассматриваемой стеганографической системе.

Ключевые слова: стеганография, стегосистема, дискретное косинусное преобразование, DCT, видеопоток.

Введение

Стеганография – это научный подход к безопасной передаче конфиденциальных данных по незащищенной сети путем их внедрения в скрывающий объект за счет наличия в нем некоторой избыточности [1][2].

Видеофайлы являются удобным и распространенным в использовании скрывающим объектом для стеганографии [3]. Так как последовательность кадров в видео имеет большую избыточность (возможность встроить большое количество данных), обусловленную большим количеством кадров в видео и особенностями восприятия человеком быстрой смены кадров, она представляет наибольший интерес для исследования. Видео в качестве скрывающего объекта может рассматриваться как файл конечного размера или как видеопоток (размер заранее не известен, внедрение данных происходит в режиме реального времени в каждый поступающий кадр).

Видеостеганография изучает две группы методов: встраивание данных в несжатое видео, которое затем сжимается, и встраивание данных непосредственно в сжатое видео. Проблема первой группы заключается в том, что получаемые стеганограммы неустойчивы к сжатию, но, несмотря на это, исследование первой группы методов является популярным направлением [3] так как сжатие видео приводит к удалению избыточных видеоданных и уменьшает пространство для сокрытия данных.

Рассматриваемая стеганографическая система

Стеганографическая система (далее – стегосистема) – это совокупность средств и методов, необходимых для формирования скрытого канала передачи данных [4].

На основе описания обязательных компонентов стегосистемы в источниках [2] и [4] и применения дискретного косинусного преобразования (DCT) в видеостеганографии в источнике [5] выделены основные компоненты стегосистемы, использующей в качестве контейнера несжатый видеопоток и основанной на методе дискретного косинусного преобразования:

1. пустой контейнер или скрывающий объект – несжатый видеопоток (последовательность поступающих в систему кадров неопределенной длины);

2. скрываемые данные – любые данные, которые могут быть представлены в виде последовательности бит;

3. стеганографический метод – метод сокрытия данных на основе дискретного косинусного преобразования;

4. стеганокодер – модуль, отвечающий за сокрытие данных в контейнере на основе заданного стеганографического метода;

5. стеганограмма – контейнер, в котором стеганокодером скрыты данные;

6. стеганодекодер – модуль, восстанавливающий скрытые данные из стеганограммы;

7. стегоканал – это канал связи, по которому передается стеганограмма.

Применяемый стеганографический метод

Дискретное косинусное преобразование (DCT) – это линейное обратимое преобразование, которое применяется к матрице (в случае с обработкой кадра матрицей является блок пикселей) для получения матрицы частотных коэффициентов, которые можно разделить на три подгруппы: низкочастотные, среднечастотные и высокочастотные [5][6]. Низкочастотные коэффициенты, расположенные в верхней левой области матрицы коэффициентов, влияют на видимое отображение блока пикселей, а высокочастотные, расположенные в нижней правой области, не влияют на видимые изменения пикселей, но сильно подвержены шумам и сжатию [7]. Для сокрытия данных подходят среднечастотные коэффициенты. Обратимость заключается в том, что из матрицы частотных коэффициентов может быть восстановлена матрица пикселей путем применения обратного дискретного косинусного преобразования (IDCT). Таким образом, DCT позволяет перейти из пространственной области изображения в частотную, что дает больше возможностей для встраивания скрываемых данных без видимого искажения кадра [6].

Далее на основе источников [5][6][7] составлена последовательность шагов сокрытия данных в видеопотоке с помощью метода на основе DCT:

1. из входного потока извлекается кадр;

2. кадр разбивается на N блоков размером 8x8 пикселей (блоки $B_i,\ i = 1,2,3 \dots, N$), пиксели в каждом блоке представляются в числовом формате;

3. к каждому B_i последовательно применяется DCT

$$C_i(u,v) = \frac{1}{4}\alpha(u)\alpha(v)\sum_{m=0}^{7}\sum_{n=0}^{7} B_i(m,n)\ cos\left[\frac{(2m+1)u\pi}{16}\right] cos\left[\frac{(2n+1)v\pi}{16}\right]$$

где $B_i(m,n)$ – элемент матрицы пикселей, расположенный на позиции (m,n), $C_i(u,v)$ – элемент матрицы частотных коэффициентов, расположенный на позиции (u,v), а $\alpha(u) = \begin{cases}\frac{1}{\sqrt{2}}, u = 0\\ 1, u > 0\end{cases}$ и $\alpha(v) = \begin{cases}\frac{1}{\sqrt{2}}, v = 0\\ 1, v > 0\end{cases}$ – нормирующие множители для нулевой частоты, в результате чего формируется матрица частотных коэффициентов также размера 8x8;

4. в полученной матрице для обрабатываемого блока изменяются коэффициенты из среднечастотного диапазона с заранее заданными координатами: выбираются пары коэффициентов $C_i(i,j)$ и $C_i(k,l)$ и, если нужно записать бит «1», коэффициенты изменяются так, чтобы выполнялось неравенство $C_i(i,j) > C_i(k,l)$, иначе, если нужно записать бит «0», то $C_i(i,j) < C_i(k,l)$. Для повышения устойчивости к помехам коэффициенты изменяются таким образом, чтобы разница между ними превышала

некоторое установленное число Q $(|C_i(i,j) - C_i(k,l)| > Q)$. Чем больше Q, тем выше надежность и ниже незаметность;

5. к массиву с измененными коэффициентами для обрабатываемого блока применяется IDCT

$$B_i(m,n) = \frac{1}{4}\sum_{u=0}^{7}\sum_{v=0}^{7}\alpha(u)\alpha(v)\,C_i(u,v)\cos\left[\frac{(2m+1)u\pi}{16}\right]\cos\left[\frac{(2n+1)v\pi}{16}\right]$$

6. полученные блоки B_i, $i = 1,2,3 \ldots, N$ объединяются в кадр;

7. собранный кадр записывается в выходной видеопоток (передается по стегоканалу).

Для извлечения скрытых данных применяется следующая последовательность шагов:

1. из стегоканала извлекается кадр;

2. кадр разбивается на блоки размером 8x8 пикселей, к которым последовательно применяется DCT аналогично сокрытию данных;

3. в полученной матрице выбираются пары коэффициентов с заранее известными координатами: $C_i(i,j)$ и $C_i(k,l)$ и, если $C_i(i,j) > C_i(k,l)$, считывается бит «1», иначе «0»;

4. считанные биты объединяются – скрываемые данные извлечены.

На основе рассмотренных состава стегосистемы и последовательности шагов сокрытия и извлечения данных, для выбранной стегосистемы разработана ее схема (рис. 1).

Рис. 1. Схема стегосистемы, использующей несжатый видеопоток и основанной на методе DCT

Сравнение метода на основе DCT с другими классическими методами

В основе скрытой передачи данных лежат три требования [8]: скрывающая способность (максимальное количество данных, которые могут быть встроены в контейнер, не влияя на его воспринимаемое качество), незаметность (степень, в которой измененный контейнер неотличим от первоначального) и надежность (способность сокрытых данных выдерживать искажения, которые могут возникнуть при хранении или передаче стеганограммы по стегоканалу).

В соответствии с предложенной в источнике [3] классификацией методов видеостеганографии среди методов, работающих с несжатыми видео, выделяют методы, действующие в пространственной области и в области преобразования. Методы, действующие в пространственной области, используют значения пикселей кадра, в которые встраивают биты секретных данных. Классические примеры таких методов: LSB, PVD и QIM [3][5][7]. Методы, действующие в области преобразования, представляют блоки кадров в виде набора коэффициентов преобразования, затем некоторые из коэффициентов изменяются (в них внедряются данные) [7]. Классические примеры таких методов: DWT и DCT [5][10]. На основе анализа источников [5][7][9][11] и представленных в них сравнений различных методов составлена таблица 1. Значения «Высокая», «Средняя» и «Низкая» по

критериям скрывающей способности, незаметности и надежности для LSB метода указаны в соответствии с выводами из источника [7], а для остальных методов – в сравнении с LSB.

Таблица 1

Сравнение классических методов видеостеганографии

Метод	Скрывающая способность	Незаметность	Надежность
LSB	Высокая	Высокая	Низкая
PVD	Средняя	Высокая	Средняя
QIM	Средняя	Средняя	Средняя
DWT	Средняя	Высокая	Высокая
DCT	Высокая	Высокая	Средняя

Достоинства и недостатки метода на основе DCT: сохраняется высокое качество кадра, высокая скрывающая способность [9], по сравнению с методами в пространственной области работает медленнее, но по сравнению с DWT реализация простая [7], стеганограмма устойчива к помехам и сжатию, но, для повышения надежности применяется помехоустойчивое кодирования (в источнике [11] упоминаются такие методы как кодирование Хэмминга и BCH-коды).

Заключение

Стегосистема, использующая в качестве контейнера несжатый видеопоток и основанная на методе дискретного косинусного преобразования, подходит для сокрытия большого количества данных. Преимуществами исследования такой стегосистемы являются простота ее реализации в сочетании с возможностью повышения надежности и высокими скрывающей способностью и незаметностью.

Так как для сокрытия данных в видеопотоке требуется быстрая обработка кадров, важным направлением для дальнейших исследований является оптимизация метода на основе DCT за счет варьирования шагов сокрытия данных в видеопотоке с учетом выполнения рассмотренных требований.

Список литературы

1. K.Sailaja, N.Madhavi, K.Sreeramamurthy, Dhanaraju Murala, K.Ashwini, Swapna Vanguru. An Efficient And Secure Data Hiding Technique: Video Steganography. Library Progress International // Library Progress International. Vol.44 No.3. Jul-Dec 2024;
2. Пузаков А.В., Аветисян К.Р. Анализ криптографических методов в целях выявления оперативно-значимой информации в стенографических контейнерах // Московский университет МВД России имени В.Я. Кикотя. 2022;
3. Anshuman Jadeja, Kush Shan, Mihan Jhaveri. Analytical Review of the Video Steganography Techniques // GIS Science Journal. Vol. 10, Issue 4. 2023;
4. Kustov V.N., Grokhotov A.I., Golovkov E.V. A Simulation Software Model of the ⊕HUGO Stegosystem // Intellectual Technologies on Transport. No 3. 2022.
5. Hammad O.M., Smaoui I., Fakhfakh A., Hashim M.M. An Overview of Image Steganography Techniques: Historical Development, Methodologies, and Evaluation Criteria // SHIFRA. Vol. 2024. 2024;
6. Чурсин Д.С. Методы внедрения контрольной информации в изображения // Научный результат. Информационные технологии. Т.9, №2. 2024;
7. Jayakanth Kunhoth, Nandhini Subramanian, Somaya Al-Maadeed, Ahmed Bouridane. Video steganography recent advances and challenges // Multimedia Tools and Applications (2023). 2023;
8. Mahmud Ahmad Bamanga, Aliyu Kamalu Babando, Mohammed Ahmed Shehu. Recent_Advances_in_Steganography // IntechOpen. 2024;

9. Shahid Rahman, Jamal Uddin, Muhammad Zakarya, Hameed Hussain, Ayaz Ali Khan, Aftab Ahmed, Muhammad Haleem. A Comprehensive Study of Digital Image Steganographic Techniques // IEEE Access. Vol. 11. 2023;
10. Munthir Bahir Tuieb, Ali Shakir Mahmood, Farah Neamah Abbas. Fusion of DWT as a Novel Approach for Efficient, Secured and Reversible Video Steganography // Journal of Physics: Conference Series 2322. 2022;
11. Wen-Bin Lin, Tai-Hung Lai, Ko-Chin Chang. Statistical feature-based steganalysis for pixel-value differencing steganography // EURASIP Journal on Advances in Signal Processing. 2021.

СЕКЦИЯ ИНФОРМАЦИОННЫЕ РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ И УСТРОЙСТВА

КРЕМНИЕВЫЕ ФОТОПРИЕМНИКИ В ЗАДАЧАХ ТЕМПЕРАТУРНОГО КОНТРОЛЯ В МЕТАЛЛУРГИИ

БАРВИНЕНКО Д. С., ЛЫСЕНКО Н. В., ФЕДОТОВ Е. В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В данной работе рассматривается возможность применения кремниевого фотоприемника в задачах температурного контроля в металлургии. Проведен анализ количества фотонов, регистрируемых фо-топриемником в ближнем инфракрасном диапазоне, для различных температур нагретых тел, характерных для доменных печей. С учетом параметров сенсора Sony IMX432 LLJ-C выполнено моделирование количества фотоэлектронов сенсора, отношения сигнал/шум при различных временах накопления и NETD. Исследование демонстрирует потенциал использования кремниевых фотоприемников для построения тепловизионных изображений в металлургических условиях без необходимости использования дорогостоящих систем со специальными фотоприемниками и охлаждением до криогенных температур

Ключевые слова: кремниевый фотоприемник, инфракрасный диапазон, квантовая эффективность, насыщение, чувствитель-ность, NETD

Контроль температуры является одной из главных задач в технологических процессах металлургической промышленности. В частности, в доменных печах, где температура может превышать 2000 ℃ точный и непрерывный контроль температуры необходим для обеспечения качества продукции, предотвращения аварийных ситуаций на производстве, контроля износа оборудования. На сегодняшний момент подавляющее большинство промышленных тепловизионных систем используют фотоприемники на основе полупроводников материалов таких как *InGaAs*, *HgCdTe*, *InSb* и т. д. Несмотря на их высокую квантовую эффективность и чувствительность в ближнем ИК диапазоне данные фотоприемники обладают рядом существенных недостатков таких как: необходимость охлаждения фотоприемника до криогенных температур ввиду экспоненциального роста темного тока, что накладывает ограничения на конструкцию самой системы, делая ее громоздкой и сложной в эксплуатации, что также сказывается на стоимости тепловизионной системы [1].

В то же время телевизионные системы, в основе которых используются кремниевые фотоприемники для работы в видимом спектральном диапазоне демонстрируют квантовую эффективность в том числе в ближней ИК области, вплоть до 1000 нм (рисунок 1), что позволяет рассматривать их в качестве альтернативы существующим тепловизионным системам [2].

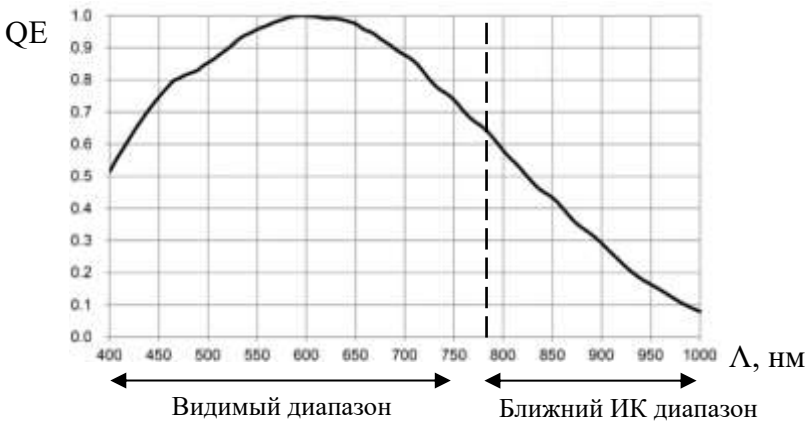


Рисунок 1 – квантовая эффективность кремниевых КМОП сенсоров Sony IMX поколения Pregius 3

Однако низкая квантовая эффективность в ближней ИК области и работа в ограниченном спектральном диапазоне накладывает определенные ограничения на температурный диапазон исследуемых объектов, модель фотоприемника и оптическую систему. В данной работе рассматривается возможность использования кремниевого фотоприемника для построения карты температур нагретых тел в температурном диапазоне специфичном для металлургической области.

Доменная печь, как сложная металлургическая установка, включает несколько характерных зон, каждая из которых имеет свои температурные особенности. Наиболее важные процессы доменной печи происходят в температурном диапазоне от 500 °C до 1400 °C таким образом целесообразно рассматривать именно его [3]. Количество фотонов в диапазоне температур 500 – 1400 °C изменяется на несколько порядков, что подразумевает работу в широком диапазоне освещенностей. Под решение данной задачи подходит фотоприемник с глобальным затвором Sony *IMX432 LLJ-C* поколения *Pregius 3*. Такие особенности как большой размер пиксела 9×9 мкм^2, глубина потенциальной ямы пиксела 99483 ē, широкий динамический диапазон 72.44 дБ, низкий шум считывания порядка 15 ē, уменьшение время накопления фотоприемника до долей времени строки делают возможным работу в широком диапазоне освещенностей.

По причине работы камеры вблизи высокотемпературных зон доменной печи предполагается установка системы в термостойкий защитный кожух с малым входным отверстием. Для обеспечения полной апертурной совместимости оптической системы с геометрией защитного кожуха без дополнительных потерь светового потока предполагается использование объектива с относительным отверстием не менее f/4.

Тогда с учетом параметров выбранного фотоприемника и оптической системы, число электронов, образованных в пикселе фотоприемника световым потоком в диапазоне длин волн [λ_1, λ_2]:

$$N_{ek}(T,\lambda)=\frac{\tau\left(\frac{D}{f}\right)^2}{4}A\cdot T_{\text{нак}}\int_{\lambda_1}^{\lambda_2}\varepsilon\frac{2\pi c}{\lambda^4}\frac{1}{e^{\frac{hc}{\lambda kT}}-1}\cdot QE(\lambda)d\lambda, \quad (1)$$

где: $T_{\text{нак}}$ – время накопления, QE(λ) – квантовая эффективность ФП на длине волны λ [4].

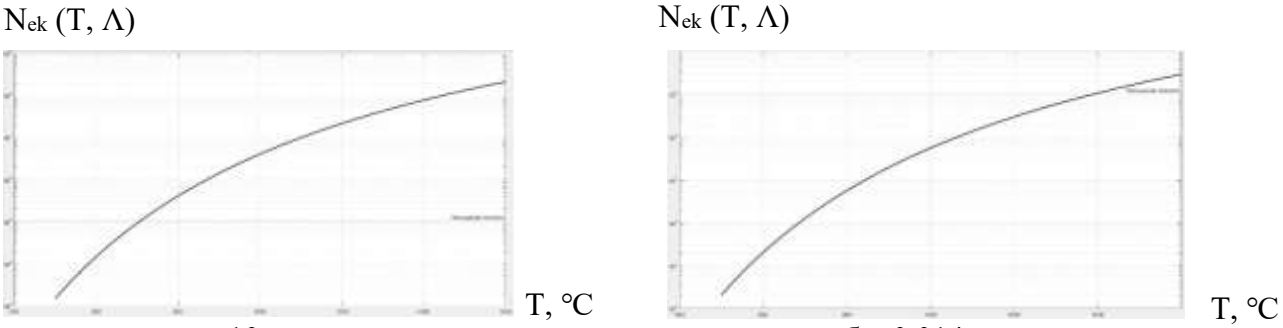
Расчеты проводились для спектрального диапазона 850…1000 нм. Ограничение в 850 нм связано с тем, что при высоких температурах наблюдаемых объектов количество фотонов, попадающих на фотоприемник, станет запредельно большим, что вызовет

деградацию видеосигнала. Верхний порог в 1000 нм был установлен по причине крайне низкой квантовой эффективности кремния в дальнейшем спектральном диапазоне.

Из графиков количества образовавшихся фотоэлектронов видно, что при максимальном времени накопления насыщение возникает при 700 ℃, уменьшая время накопления до 0,014 мс насыщения пиксела сдвигается до температуры в 1400 ℃ (рисунок 2). Для оценки качества получившихся тепловизионных изображений при различных временах накопления необходимо провести оценку отношения сигнал/шум. При появлении сигнальных электронов возникает составляющая шума, называемая фотонным шумом, то есть дробовой шум светового потока. Фотонный шум подчиняется распределению Пуассона, в котором матожидание равно дисперсии, тогда отношение сигнал/шум:

$$\Psi = \frac{N_{ek}}{\sqrt{N_{ek} + N_{\text{шсч}}{}^2}}, \tag{3}$$

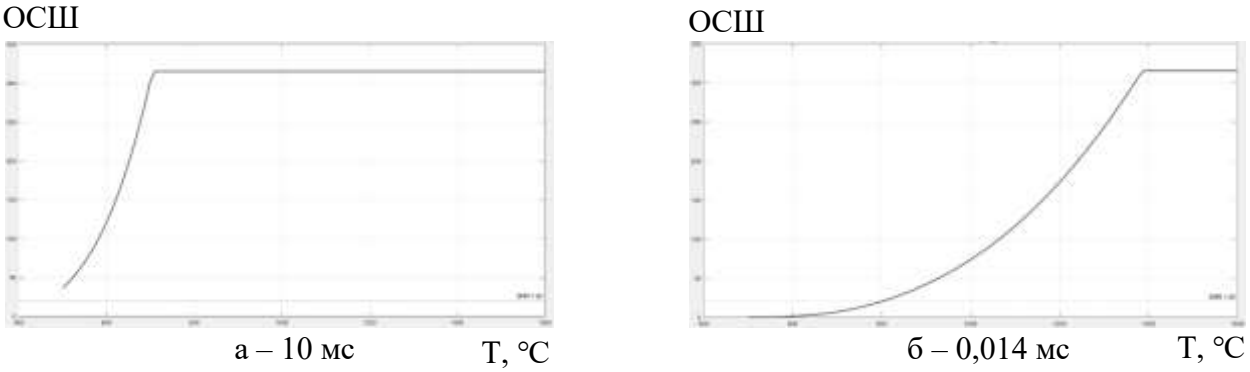
где $N_{\text{шсч}}$ – шум считывания, равный 15 ē.



а – 10 мс б – 0,014 мс

Рисунок 2 – Зависимость количества образовавшихся фотоэлектронов в пикселе фотоприемника от температуры для различных времен накопления

Рисунок 3 показывает рост ОСШ в зависимости от температуры для максимального времени накопления. Минимальное значение ОСШ соответствует значению 36. Наилучшее значение ОСШ достигается при максимальном зарядовом пакете. В то же время при малых временах накопления, при наблюдении слабо нагретых тел ОСШ не удовлетворительное ввиду малого значения сигнала. Таким образом для получения оптимального ОСШ необходима программная корректировка времени накопления фотоприемника в режиме реального времени при работе тепловизионной камеры.



а – 10 мс б – 0,014 мс

Рисунок 3 – ОСШ при различных временах накопления

Температурная чувствительность является важнейшим параметром тепловизионной системы. В данной работе рассматриваются два подхода: теоретическое температурной разрешение и NETD (noise equivalent temperature difference – разность температур эквивалентная шуму). Температурное разрешение определяется как минимальное изменение температуры, которое приводит к изменению выходного сигнала на один уровень АЦП. Шаг квантовая фотоприемника Sony IMX432 LLJ-C составляет $\Delta Q=99483/2^{12}=24{,}3$ электрона на один уровень АЦП. Изменение выходного сигнала фотоприемника при изменении температуры исследуемого объекта характеризуется производной dN_e/dT. Тогда теоретическое температурное разрешение возможно выразить как:

$$\triangle T_{quant} = \frac{\triangle Q}{\frac{dN_e}{dT}}$$

Данный параметр показывает теоретический предел чувствительности фотоприемника Sony IMX432LLJ-C при отсутствии шумов, исключительно ограниченный разрядностью АЦП и глубиной потенциальном ямы пиксела. Однако на практике всегда присутствует такие составляющие шума как фотонный шум и шум считывания. Стоит отметить, что шум темнового тока для кремниевых фотоприемников незначителен, его экспоненциальный рост начинается при высоких рабочих температурах фотоприемника, соответственно в расчётах он приравнивается к нулю [5]. Параметр NETD рассчитывается по выражению:

$$NETD = \frac{\sqrt{N_e + \sigma_{read}^2}}{\frac{dN_e}{dT}},$$

NETD показывает реальную чувствительность системы, ограниченную шумами и используется как основной метрический параметр в технических спецификациях тепловизоров.

Температурная чувствительность

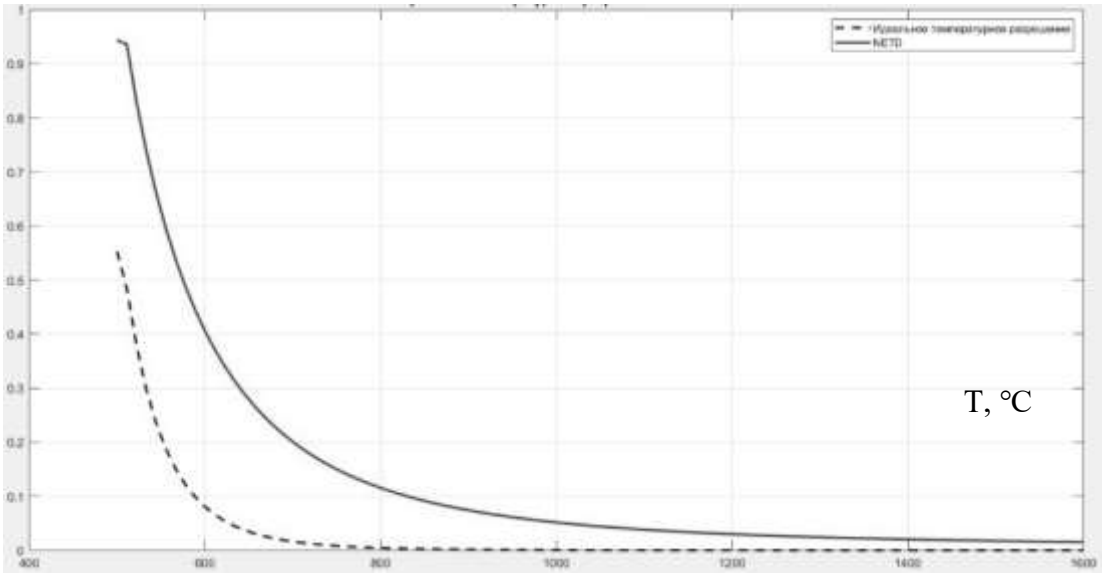


Рисунок 4 – Сравнение предела температурного разрешения Sony IMX432LLJ-C и NETD для времени накопления 0,01 с

Таким образом в работе показана эффективность использования кремниевого фотоприемника для построения карты температур нагретых тел в диапазоне 500…1400 °C. При времени накопления 0.01 с фотоприемник способен обеспечивать температурное разрешение (NETD) порядка 0.2 °C, при этом достигается максимальное отношение сигнал/шум около 315. Это подтверждает возможность применения кремниевых фотоприемников в металлургии при корректировке чувствительности сенсора для предотвращения насыщения и обеспечения высокой точности измерений.

Список литературы

1. Rogalski A. Infrared detectors: status and trends // Progress in Quantum Electronics. 2003. Vol. 27, iss. 2–3. P. 59–210. DOI: 10.1016/S0079-6727(02) 00024-1.
2. https://www.npk-photonica.ru/product/1213/
3. Вегман Е. Ф., Жеребин Б. Н., Похвиснев А. Н. и др. Металлургия чугуна : Учебник для вузов / под ред. Ю. С. Юсфина. — 3-е издание, переработанное и дополненное. — М. : ИКЦ «Академкнига», 2004. — 774 с.
4. Манцветов А. А., Барвиненко Д. С., Ульяновский К. А., Баранов П. С. Оценка уровня сигнала высокотемпературном тепловизионной камеры// Вопросы радиоэлектроники. Сер. Техника телевидения 2024. Вып. 6.
5. Анализ количества темновых электронов твердотельных фотоприемников при высокой рабочей температуре / К. В. Стрижнев, Д. А. Белоус, П. С. Баранов, В. Т. Литвин, А. А. Манцветов, В. А. Михайлов // Вопросы радиоэлектроники. Сер. Техника телевидения. 2017. Вып. 2. С. 31–40.

ОРГАНИЗАЦИЯ ОБМЕНА ДАННЫМИ В УСТРОЙСТВЕ РЕГИСТРАЦИИ СИГНАЛОВ АКУСТИЧЕСКОГО ДАТЧИКА

В.А. БЕСПАЛОВ, Г.В. БУЗУВЕРОВ, О.М. АНДРЕЕВА

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В работе представлено описание решения задачи организации обмена данными между блоками макета устройства регистрации акустических сигналов. Описано три линии связи: АЦП-Raspberry Pi, GPS-модуль-Raspberry Pi, Raspberry Pi-LoRa.

Ключевые слова: микрокомпьютер, обмен данными, программный продукт, АЦП, GPS, LoRa.

Введение

Акустический мониторинг – это способ контроля за состоянием объектов, процессов или окружающей среды при помощи звуковых волн. Данный вид наблюдения применяется в промышленности [1], экологии [2], медицине [3].

Создание устройства акустического контроля – многоэтапный процесс, требующий возможности выполнения многократных испытаний. Поэтому, как правило, на первоначальном этапе создается макет устройства, позволяющий провести всестороннее тестирование, убедиться в правильности работы всех блоков и корректном обмене данными между ними, отладить алгоритм обработки сигнала.

Разрабатываемое устройство регистрации сигналов акустического датчика должно выполнять обработку поступившей информации и беспроводную передачу результатов обработки. Структурная схема макета данного устройства приведена на рисунке 1.

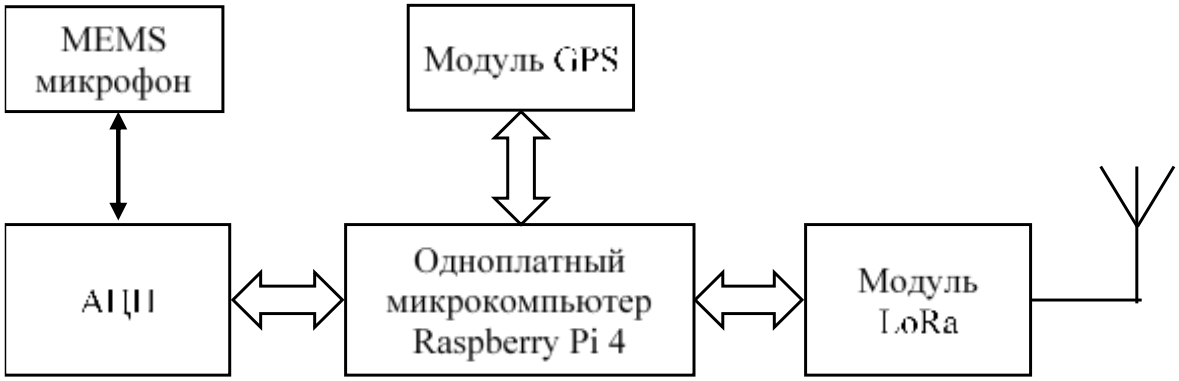


Рис. 1 Структурная схема макета

Оцифрованный акустический сигнал и информация о реальном времени, получаемая из модуля GPS, поступают на микрокомпьютер Raspberry Pi 4, в котором производится их обработка с помощью ранее созданного программного обеспечения (ПО), основанного на алгоритме машинного обучения. Результат работы ПО – классификация принятой информации и время прихода сигнала (в случае его наличия) – должен быть передан с помощью модуля LoRa в центр сбора информации.

В рамках данной работы была организована передача данных между составными элементами макета. Ниже поэтапно рассматривается решение данной задачи.

Согласование работы Raspberry Pi и АЦП

Передача данных между АЦП и микрокомпьютером производится с использованием интерфейса SPI [4], в качестве АЦП используется модуль ADC1263 [4]. Особенностью алгоритма обработки данных является требование жестко структурированного формата входной информации, а именно массива отсчетов в десятичной системе счисления и временной метки поступления последнего отсчета.

Кроме того, процесс сбора и формирования данных для ПО требует параллельного выполнения обработки ранее поступившей информации, так как нарушение данного условия приведет к потере данных и невозможности работы системы в реальном времени.

Производитель предоставил ПО на Python и C, обеспечивающее настройку АЦП, взятие отсчетов и преобразование их в десятичный формат [5]. Изучение данного ПО выявило следующие недостатки:

- код на Python имеет большие неоправданные вычислительные затраты и ограничения по производительности;
- код на C громоздкий и сложный, также имеет высокие вычислительные затраты, несовместим с данным Raspberry Pi и требует сопряжения с Python-окружением;
- отсутствие поддержки параллельных вычислений.

Поэтому было принято решение создать своё ПО настройки и управления АЦП, а также формирования данных структурированного формата.

Разработанный алгоритм формирования данных представлен на рисунке 2 и условно может быть разделен на три параллельно работающих потока:

Рис. 2 Алгоритм формирования данных

1) поток сбора данных: группирует поступающие в шестнадцатеричном формате отсчеты; фиксирует временную метку поступления последнего отсчета пачки и записывает ее в пакет данных наряду с самой пачкой. Сформированный пакет данных передается в промежуточный буфер, после чего начинается формирование нового пакета данных;

2) поток подготовки данных: ожидает наличия данных, преобразует данные в десятичный формат, сохраняет в структуру;

3) поток транспортировки: предоставление Python-окружению ПО обработки данных структуры с массивом данных и меткой времени.

Реализация алгоритма была осуществлена на языке C++ с дальнейшей интеграцией в Python-окружение.Реализация алгоритма была осуществлена на языке C++ с дальнейшей интеграцией в Python-окружение.

Для проведения верификации работы созданного ПО на блок АЦП с генератора аналоговых сигналов подавались тестовые сигналы (гармонический, прямоугольный и треугольный) с последующим анализом в Matlab в частотной и временной областях. Анализ показал корректную оцифровку сигналов, передачу временных меток и большее быстродействие созданного ПО по сравнению с ПО производителя.

Согласование работы Raspberry Pi и модуля GPS

В состав макета включен UART GPS модуль Waveshare NEO-M8T**.**

Данная модель GPS предоставляет источник высокоточного сигнала PPS – периодической последовательности прямоугольных видеоимпульсов, с периодом, равным с высокой точностью 1 секунде. Этот сигнал использовался для высокоточной синхронизации. Выбрана допустимая погрешность времени – 10^{-6} секунды.

После анализа протоколов временной синхронизации, анализа системных файлов ОС Raspberry Pi, изучения модуля GPS Waveshare NEO-M8T и его функциональных возможностей, были выбраны следующие программные инструменты [6] настройки обмена данными:

1) Фоновый процесс обработки данных с GPS-приемника gpsd;

2) Высокоточный NTP-клиент chrony.

Для интеграции модуля использовался системный конфигурационный файл config.txt платформы Raspberry Pi, в котором был выделен отдельный GPIO-вывод для сигнала PPS, а также активирован UART порт.

Проведена верификация, позволяющая утверждать корректность настройки системных файлов и согласованность работы модулей.

Согласование работы Raspberry Pi и модуля LoRa

Результат работы ПО на Raspberry Pi, которым является классификация и время прихода поступившего сигнала, должен быть передан посредством беспроводной связи далее во внешнюю среду. В качестве передатчика используется модуль LoRa. Для организации взаимодействия между микрокомпьютером и модулем было изучено руководство пользователя [7], после чего разработан и реализован программно-аппаратный протокол обмена данными.

Реализованный протокол обмена включает в себя следующий алгоритм настройки и организации передачи пакета данных:

1) Аппаратными методами производится установка модуля в режим инициализации;

2) Осуществляется посылка байтов инициализации устройства (настройка адреса устройства и канала, выбор скорости работы и передачи данных), с последующей проверкой загруженных байтов;

3) Выполняется аппаратная активация режима передачи данных;

4) Производится формирование пакета данных для модуля LoRa, отправка сформированного пакета в модуль LoRa.

В рамках верификации работоспособности протокола передачи данных проведена серия лабораторных испытаний, включающая передачу априорно известных данных. На приемной стороне осуществлялся сравнительный анализ принятых данных с исходной последовательностью. Результаты экспериментального исследования подтвердили соответствие излученного сигнала входным данным, а также отсутствие ошибок в принятом информационном потоке.

Заключение

Реализованы взаимодействия между блоками макета, проведено испытание макета в лабораторных условиях с использованием тестового ПО обработки информации. В рамках дальнейшей работы планируется проведение полигонных испытаний.

Список литературы

1. Соколов, И.М. Фракталы [Текст] / И.М. Соколов // Квант. – 1989. - №5. – С. 6-13.
2. Епифанов, Е. Красота повтора [Текст] / Е. Епифанов // Популярная механика. – 2009. - №3.
3. Слюсар, В. Фрактальные антенны принципиально новый тип "ломаных" антенн [Текст] / В. Слюсар // ЭЛЕКТРОНИКА: Наука, Технология, Бизнес. – 2007. – С. 78-83.
4. Фрактальные антенны [Электронный ресурс] / URL: https://bloganten.ru/fraktalnye-antenny
5. Фрактальная wi-fi антенна [Электронный ресурс] / URL: https://bloganten.ru/fraktalnaya-wifi-antenna

ЧАСТОТНАЯ СЕЛЕКЦИЯ ДВИЖУЩИХСЯ ЦЕЛЕЙ В РАДИОЛОКАЦИОННЫХ СИСТЕМАХ СО СЛОЖНЫМ СИГНАЛОМ

Зинченко А.А.

Новгородский государственный университет имени Ярослава Мудрого

Аннотация. В статье описывается метод селекции движущихся целей на основе частотной режекции помех в радиолокационных системах со сложным сигналом, обладающим двухуровневой нерегулярной структурой огибающей энергетического спектра. Приведен алгоритм формирования сложных сигналов с указанной спецификой спектра. На примере продемонстрировано, насколько перекрываются спектры сигналов с разными смещениями по частоте, а также показана трансформация спектров после частотной режекции.

Ключевые слова: селекция движущихся целей, сложные сигналы, двухуровневая огибающая амплитудного спектра, частотная режекция

Введение

Селекция движущихся целей основана на различии в смещениях по частоте спектров сигналов, отраженных от движущихся и неподвижных целей. Традиционно для решения задач обнаружения движущихся целей в радиолокационных системах (РЛС) применяют зондирующие сигналы в виде пачки последовательности периодических импульсов. Амплитудный спектр сигнала обладает гребенчатой структурой, что позволяет выполнить задачу различения спектров доплеровских целей. Однако многопиковая структура ограничивает диапазон доплеровских частот и приводит к появлению «слепых» элементов дальности.

Проблема частотной селекции движущихся целей в РЛС со сложными сигналами состоит в полном перекрытии спектров отраженных сигналов. Оригинальная идея селекции движущихся целей на фоне помех на основе формирования сложных сигналов с двухуровневой нерегулярной структурой огибающей амплитудного спектра была предложена в работах [1-4].

В настоящей статье излагается метод синтеза сложных импульсных сигналов с указанными спектральными свойствами. Дается оценки степени подавления помехи на основе частотной режекции спектральных компонент в спектре сигнала помехи.

Формирование последовательности с двухуровневой структурой огибающей амплитудного спектра

В основе формирования сигнала с нерегулярной (псевдослучайной) структурой огибающей амплитудного спектра лежит явление частотно-временного дуализма. Пусть модуль некоторой троичной последовательности, периодическая автокорреляционная функция которой имеет практически равномерный уровень боковых лепестков во всем диапазоне временных сдвигов, будет описывать форму спектра формируемого сигнала. Тогда ее фазовый спектр будет описывать закон многопозиционной фазовой манипуляции сигнала, обладающего нерегулярной структурой огибающей энергетического спектра. Соответствующие дискретные значения фазы мы можем получить из аргумента дискретного преобразования Фурье (ДПФ), представленного формулой:

$$W_k := \frac{1}{N} \cdot \sum_{i=0}^{N-1} \left(w_i \cdot \exp\left(-i \cdot 2 \cdot \pi \cdot \frac{k}{N} \cdot i \right) \right),$$

где w_i – дискретные отсчёты порождающей троичной последовательности.

Пример порождающей троичной последовательности со средней скважностью Q = 4 представлен на рисунке 1. Закон изменения фазы многофазной последовательности с непрерывной огибающей и амплитудный спектр с нерегулярной двухуровневой структурой представлены на рисунках 2 и 3 соответственно. Следует отметить, что приведенный амплитудный спектр обладает двухуровневой структурой. Выделяются спектральные компоненты с высоким уровнем интенсивности и с низким (нулевым) уровнем. Частотный интервал следования компонент с высоким и низким уровнем имеет нерегулярный характер.

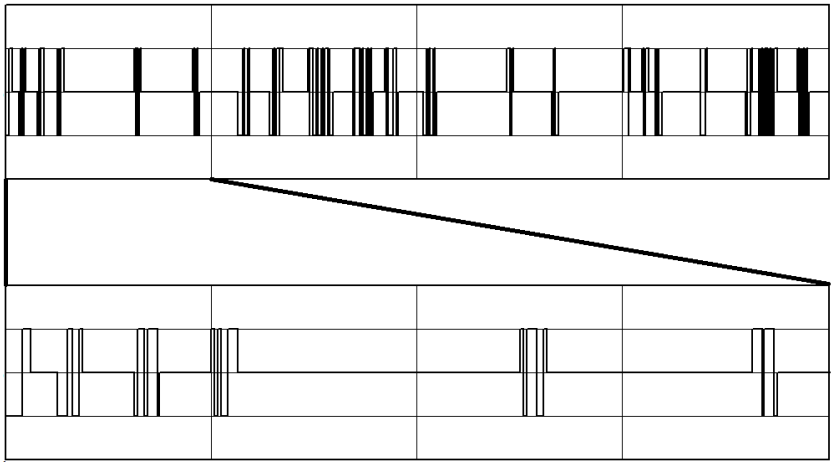


Рис. 1. Порождающая троичная псевдослучайная последовательность

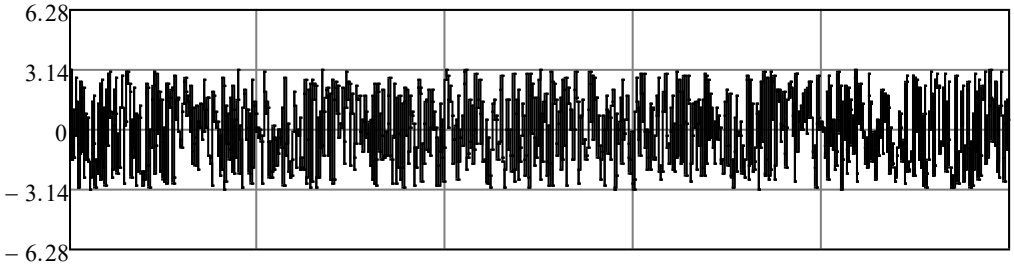


Рис. 2. Временная диаграмма изменения фазы сигнала

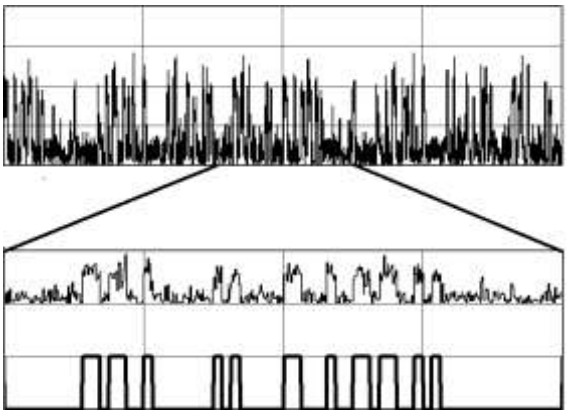


Рис. 3. Амплитудный спектр сигнала и сопоставление части спектра с модулем соответствующей части троичной порождающей последовательности

Частотная режекция мешающего точечного отражения

Степень подавления помехи определяется частотной функцией режекции. Как отмечалось ранее, двухуровневый вид амплитудного спектра однозначно определяется структурой троичной последовательности. Поэтому вид частотной функции режекции получается «инверсией» символов модуля троичной последовательности и принимает два уровня «0» или «1».

В силу того, что принимаемый сигнал является аддитивной, линейной смесью помехи и сигнала, то можно рассматривать раздельно, как изменяются спектры помехи и сигнала с различными доплеровскими сдвигами при частотной режекции.

Рассмотрим, как изменяются спектры сигналов с различными частотными сдвигами после частотной режекции последовательностью, инверсной модулю порождающей последовательности. Спектры до и после режекции представлены на рисунке 4 для сигнала с нулевым доплеровским сдвигом и на рисунке 5 для сигнала со смещением по частоте.

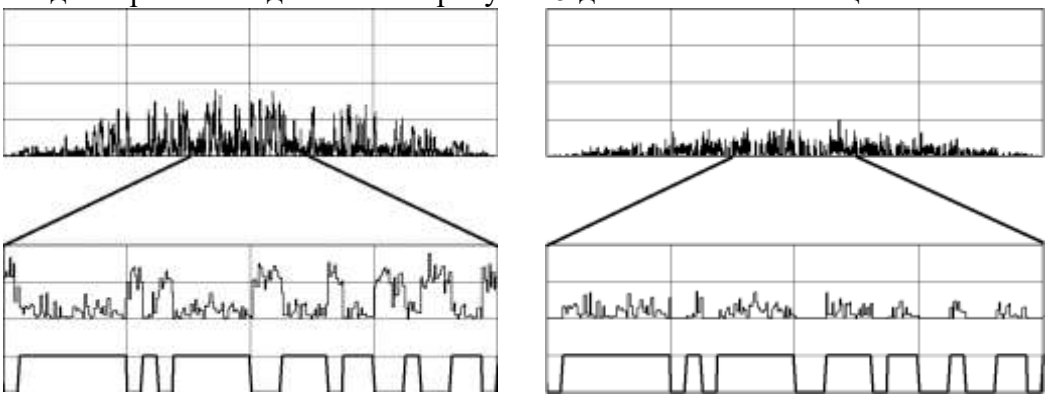


Рис. 4. Амплитудный спектр сигнала с нулевым доплеровским сдвигом до режекции (слева) и после (справа)

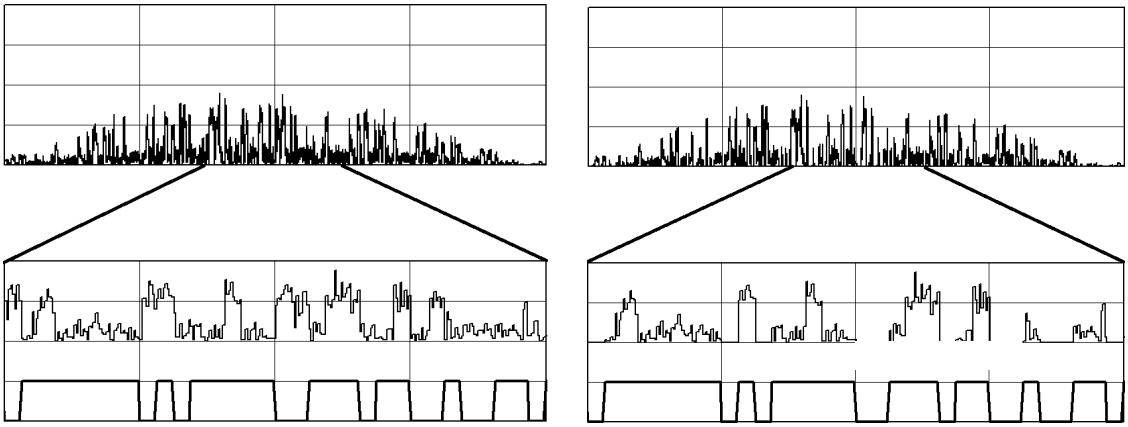


Рис. 5. Амплитудный спектр сигнала с доплеровским сдвигом до режекции (слева) и после (справа)

Из рисунка 4 видно, что высокоуровневые составляющих спектра сигнала с нулевым доплеровским сдвигом в ходе частотной режекции подавляются. Однако составляющие с низким уровнем сохранились. В результате режекции имеем подавление сигнала с нулевым доплеровским сдвигом на 16,3 дБ. Рисунок 5 демонстрирует, что спектр сигнала, имеющий отличный от нуля доплеровский сдвиг, после частотной режекции сохраняет нерегулярную двухуровневую структуру огибающей спектра. Однако, в ходе режекции теряется малая часть высокоуровневых спектральных компонент. Доля энергетических потерь при этом будет зависеть от средней скважности порождающей последовательности. Таким образом,

когерентное накопление сигналов после частотной режекции будет изменять отношение сигнал-помеха в сторону увеличения.

Следует отметить, что после проведения селекции двухуровневый нерегулярный спектр сохраняет единственность пика функции неопределенности и разрешающие способности по задержке и частоте остаются такими же, как и до режекции. Поэтому после подавления помехи во входной смеси сигнала с помехой указанным методом производиться классическая обработка сигнала, обеспечивающая максимизацию отношения сигнал/шум и решению задач радиолокационного обнаружения.

Выводы

Исследование трансформации спектра сигнала с двухуровневой нерегулярной структурой энергетического спектра в ходе частотной режекции продемонстрировало целесообразность метода для решения задач селекции движущихся целей.

Список литературы

1. Акустические методы контроля и диагностики [Электронный ресурс]. URL: https://portal.tpu.ru/SHARED/l/LEX-K/lessons/Vibro/Tab3/Акустические%20методы%20контроля.pdf
2. Акустическая экология [Электронный ресурс]. URL: file:///C:/Users/Владимир/Downloads/Акустическая%20экология.pdf
3. Физические основы использования ультразвука в медицине [Электронный ресурс]. URL: https://rsmu.ru/fileadmin/templates/DOC/Faculties/PF/Phys-mat/uzi.pdf
4. High-Precision AD HAT ADS1263 [Электронный ресурс]. URL: https://www.waveshare.com/wiki/High-Precision_AD_HAT#Python_2
5. Github_High-Pricision_AD_HAT [Электронный ресурс]. URL: https://github.com/waveshare/High-Pricision_AD_HAT
6. GNSS setup [Электронный ресурс]. URL: https://kaitotek.com/resources/documentation/probe/install/debian-ubuntu/gps-setup#step_by_step_instructions_for_chrony
7. E32-433T30D User Manual [Электронный ресурс]. URL: https://voltiq.ru/datasheets/ebyte/E32-433T30D-user-manual.pdf

МЕТОД КОМПЕНСАЦИИ ХРОМАТИЧЕСКИХ АБЕРРАЦИЙ ПРИ ФОРМИРОВАНИИ ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ, ПОЛУЧЕННЫХ С ПОМОЩЬЮ ПЕРЕСТРАИВАЕМОГО ИСТОЧНИКА СВЕТА

ГОНОШИЛКИН Д. Ю., БАРАНОВ П. С.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Аннотация. Представлен метод количественной оценки и коррекции хроматических аберраций в гиперспектральных системах, использующих последовательное спектральное сканирование и перестраиваемый LED-осветитель. Предлагается двухэтапная схема: глобальное выравнивание каналов на основе матрицы гомографии и уточнение плотным оптическим потоком Фарнебека. Проведён эксперимент на кубе 400–1000 нм с 17 каналами; остаточная RMS-ошибка совмещения снижена с 0,39 до 0,04 px. Метод реализуется в Python и не требует аппаратных модификаций системы.

Ключевые слова: гиперспектральное изображение, хроматические аберрации, гомография, оптический поток.

Введение

Гиперспектральные камеры со спектральным сканированием позволяют регистрировать пространственно-спектральный куб высокого разрешения; однако при перестройке длины волны возникают хроматические аберрации, вызывающие субпиксельные смещения изображений. Если их не компенсировать, совмещённость каналов нарушается и спектральные векторы теряют достоверность. В литературе предложены как аппаратные, так и программные способы уменьшения данного эффекта [1]–[3]. В работе рассматривается программная коррекция, не требующая изменения оптики.

Оценка хроматических аберраций

Для количественной оценки латеральной составляющей аберрации применяется контрольная сетка 22 × 14 круглых маркеров, наносимая в плоскости объекта. Координаты центров маркеров извлекаются субпиксельно на каждом канале; в качестве эталона выбирается средний по спектру кадр ($\lambda \approx 700$ нм). Вектор смещения $\Delta x, \Delta y$ между соответствующими центрами вычисляется для всех каналов; среднеквадратическое отклонение RMS служит интегральной метрикой ошибки совмещения.

До коррекции максимальная ошибка наблюдалась на периферии кадра и достигала 0,58 px при среднем значении 0,39 px.

Гомографическое выравнивание

Гомография описывает глобальное проективное преобразование, аппроксимирующее межканальные смещения. Предполагается, что аберрации однородны по полю зрения и могут быть устранены единой матрицей преобразования $\boldsymbol{H}$.

На каждом изображении локализуются опорные точки (центры маркеров калибровочной сетки), и по их соответствиям между каналом и эталоном с помощью метода RANSAC рассчитывается матрица $\boldsymbol{H}_{\lambda}$. После этого канал трансформируется в координатную систему эталона, сохраняя пространственное разрешение.

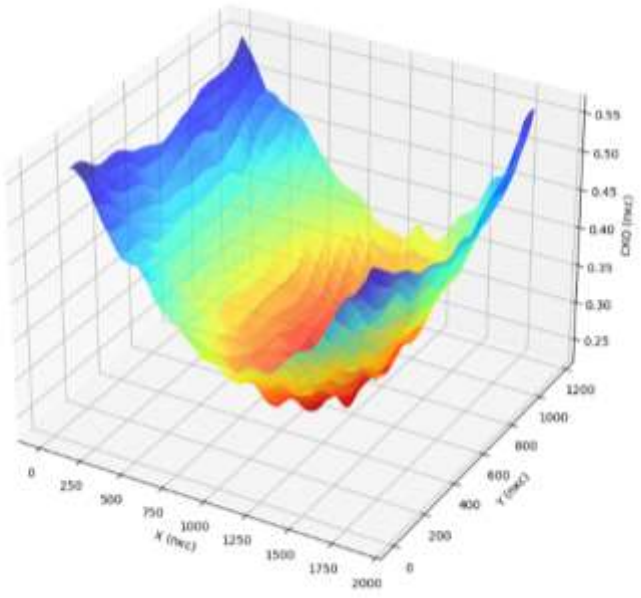


Рис. 1. Распределение среднего смещения от центрального канала на изображении до коррекции

Оптический поток Фарнебека

Метод оптического потока позволяет вычислять плотное поле смещений между каналами, оценивая локальные изменения яркости. В данной работе используется алгоритм Фарнебека, обеспечивающий субпиксельную точность и устойчивость к слабому текстурному контрасту [4].

Каждый канал выравнивается относительно эталонного с помощью построенного векторного поля $V(x, y) = (u, v)$. При этом используется интерполяция для сохранения структуры изображения.

Данный подход более универсален и позволяет учитывать сложные и радиально-зависимые аберрации, но требует больше вычислительных ресурсов.

Экспериментальные результаты

Метод протестирован на гиперспектральной системе с перестраиваемым LED-осветителем (400–1000 нм, 17 спектральных каналов). Совмещение оценивалось по 308 меткам сетки.

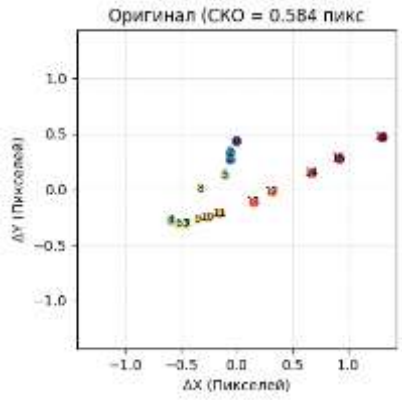
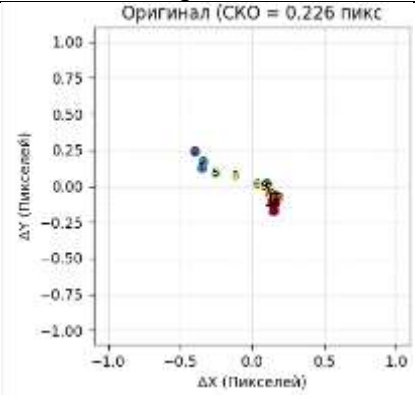
Таблица 1 Сводные значения RMS-ошибки.

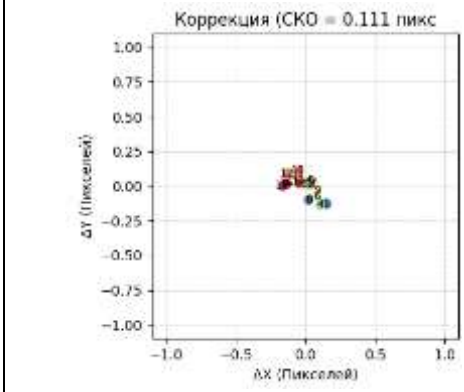
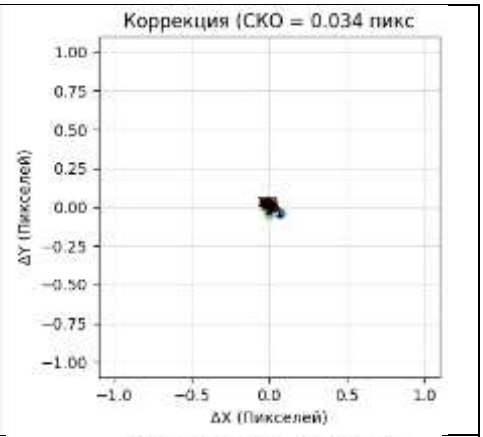
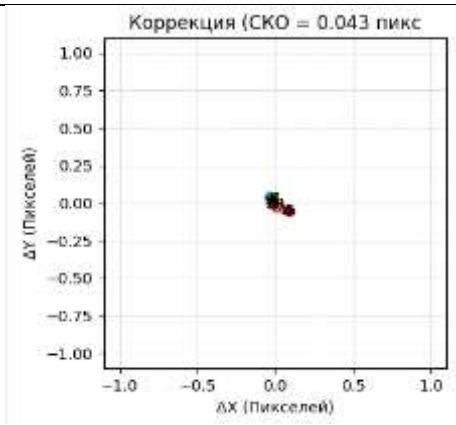
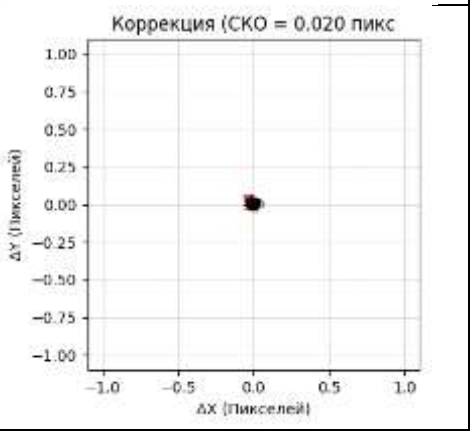
Результаты сравнения методов коррекции хроматической аберрации

Метод	RMS, px средн.	RMS, px макс.	Время, с
До коррекции	0,39	0,58	—
Гомография	0,11	0,24	0,12
Оптический поток	0,04	0,07	0,40

Полученные результаты подтверждают эффективность комбинированного подхода: интегральная ошибка уменьшена в 9 раз без вмешательства в аппаратную часть комплекса.

Таблица 2 Результаты проведенной коррекции

Используемое изображение	Исследуемая точка	
	Наиболее искаженная точка изображения	Наименее искаженная точка изображения
Изображение до коррекции		

Изображение после гомографической коррекции		
Изображение после коррекции оптическим потоком		

Заключение

Рассмотрены методы компенсации хроматических аберраций, с помощью гомографической коррекции и с использованием оптического потока. Экспериментально показано, что они обеспечивает субпиксельную точность и могут быть использованы как этап предпроцессинга в системах спектрального сканирования с перестраиваемым источником света. при приоритетной задаче минимизации межканальных смещений оптический поток обеспечивает лучшую геометрическую точность. Гомография рекомендована как быстрая «первичная» корректировка либо как резервный вариант для систем с крайне ограниченными ресурсами обработки. В дальнейшем планируется улучшения алгоритма для возможности корректирования смещения фокуса вследствие хроматической аберрации.

Список литературы

1. Mallon J., Whelan P. F. Calibration and removal of lateral chromatic aberration in images. Pattern Recognition Letters, 28(1), 2007.
2. González-Hernández J. A., et al. Characterization and correction of lateral chromatic aberration in multichannel imaging systems. Optics & Lasers in Engineering, 163, 2023.
3. Yokoya N., et al. Cross-sensor image matching for hyperspectral images using optical-flow estimation. ISPRS J. Photogrammetry & Remote Sensing, 145, 2018.
4. Ghamisi P., Yokoya N. Hyperspectral image alignment employing subpixel optical flow. IEEE Geoscience and Remote Sensing Letters, 17(8), 2020.

ИССЛЕДОВАНИЕ СТРУКТУРЫ СИГНАЛОВ С ПОМОЩЬЮ ПРЕОБРАЗОВАНИЙ ВЕЙВЛЕТОВ, ВИГНЕРА-ВИЛЯ И РАДОНА

ГАЙВОРОНСКИЙ Д. В., ДИАНОВ А. В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В работе представлены результаты исследования методов анализа структуры сигналов, альтернативных классическому преобразованию Фурье и основанных на преобразованиях вейвлетов, Вигнера-Виля и Радона. На основе используемых преобразований синтезируются новые методы оценки таких параметров сигналов, как время запаздывания и доплеровское смещение частоты.

Ключевые слова: вейвлет-преобразование, преобразование Вигнера-Виля, преобразование Радона, скейлограмма, простые и сложные сигналы, оценка времени запаздывания, доплеровское смещение частоты.

Начиная с восьмидесятых годов двадцатого века большими темпами идёт развитие теории вейвлет-преобразования, являющейся перспективным и удобным математическим аппаратом, практические методы использования которого позволяют получить ответы на интересующие вопросы, например, в такой области, как цифровая обработки изображений, применительно к анализу влияния мелких деталей, сжатия и прочее. В ходе проведённых исследований рассматривались вопросы применения математического аппарата вейвлет-преобразования к таким радиотехническим задачам, как разложения сигнала в ряд с помощью базисных вейвлет-функций, а также получение информации о ключевых параметрах сигналов [1], [2].

Возможность решения этих задач обусловлена самой сущностью вейвлета, который, вне зависимости от принадлежности к конкретному семейству, за счёт конечной протяжённости вдоль оси времени и частот позволяет наблюдать характерные особенности в поведении сигнала, как во временной, так и в частотной областях (см. рис. 1 и 2).

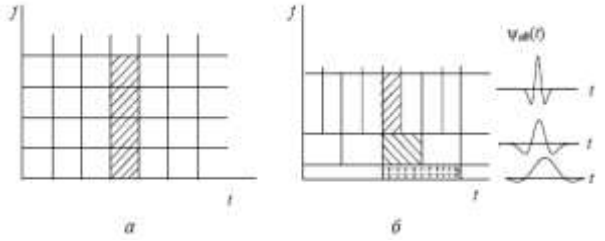


Рис. 1

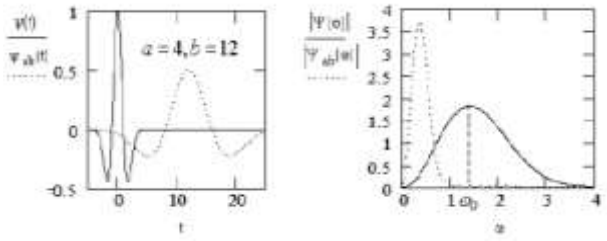


Рис. 2

Изучение теоретического материала и проведённые исследования показали, что наиболее важным с точки зрения решения поставленной задачи оценки параметров сигналов является инструмент прямого непрерывного вейвлет-преобразования (ПНВП):

$$W_S(a,b) = \left(S(t), \psi_{ab}(t)\right) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} S(t)\psi\left(\frac{t-b}{a}\right) dt.$$

Осуществляя прямое непрерывное вейвлет-преобразование при нескольких разных значениях параметров a и b можно получить частотно-временное представление сигнала (параметр b характеризует временные изменения сигнала, а параметр a – частотные), которое называется скейлограммой, и произвести оценку интересующих параметров сигналов.

Другими методами получения информации об изменениях параметров сигнала, аналогичными вейвлет-преобразованию, являются преобразования Вигнера-Виля и Радона.

Преобразование Вигнера-Виля принадлежит к классу совместных время-частотных распределений и имеет множество сходств с характеристической функцией неопределённости Вудворда [3]:

$$P(t,\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} s(t+\tau/2)\, s^*(t+\tau/2) \exp(-j\omega\tau)\, d\tau.$$

Визуализация результатов преобразования идейно очень похожа на скейлограмму и представляет собой двумерный график линий уровня на частотно-временной плоскости. На рисунках 3 и 4 представлены результаты преобразования Вигнера-Виля над прямоугольным радиоимпульсом и сигналом с линейной-частотной модуляцией, соответственно.

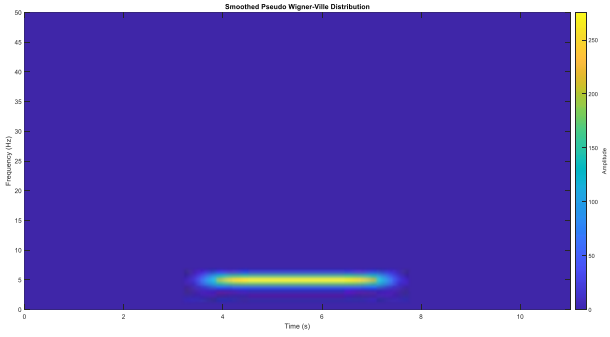


Рис. 3

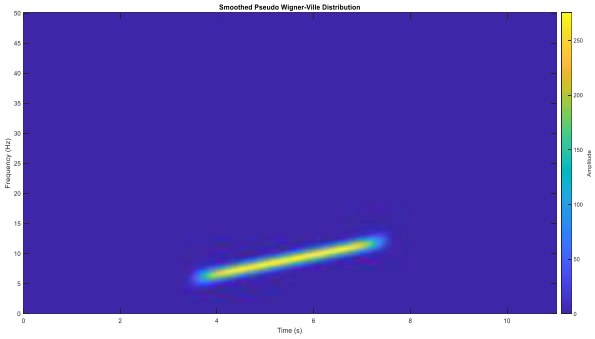


Рис. 4

Метод Радона представляет собой преобразование, родственное преобразованию Фурье, которое вводится с помощью следующего выражения:

$$R(s,\alpha) = \int_{L=AA'} f(x,y)dL = \int_{-\infty}^{+\infty} f(s\cdot\cos(\alpha) - z\cdot\sin(\alpha)\,, s\cdot\sin(\alpha) + z\cdot\cos(\alpha))dz.$$

В проведённых исследованиях в качестве функции $f(x, y)$ рассматривалась поверхность, сформированная с помощью набора сигналов, построенных на заданном временном промежутке, каждый из которых имеет свою несущую частоту. С помощью преобразования Радона появляется возможность сделать оценку частотных и временных параметров сигнала в соответствии с ходом прямой AA' (см. рис. 5 и 6).

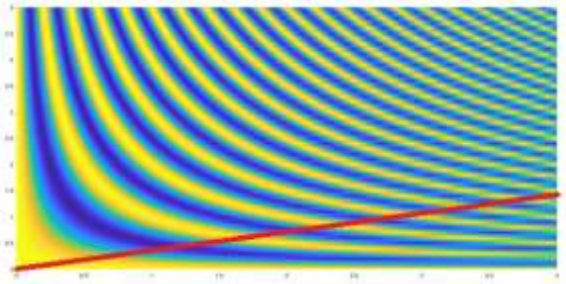


Рис. 5

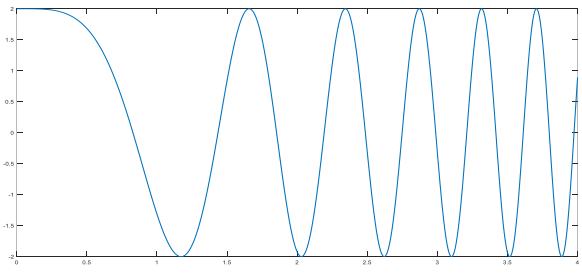


Рис. 6

Предполагается, что рассмотренные преобразования будут заложены в основу методов оценки параметров сигналов, альтернативных классическим методам, реализованных с

помощью структуры из согласованных фильтров или корреляторов [4]. При этом оцениваемыми параметрами являются время запаздывания и доплеровское смещение частоты сигнала.

Предлагаемый метод оценки времени запаздывания предполагает, что программным образом осуществляется вычитание вейвлет-преобразование принятого сигнала и вейвлет-преобразование принятого сигнала задержанного или опережающего на некоторые моменты времени. При этом минимальная разность будет достигаться в том случае, если искусственно внесённый временной сдвиг совпадает с истинным значением времени запаздывания. Структурная схема, осуществляющая такую обработку, является многоканальной и при переходе от одного канала к другому значение искусственно вносимого временного сдвига меняется.

Метод оценки доплеровского смещения частоты предполагает, что программным образом осуществляется вычитание вейвлет-преобразования принятого сигнала и вейвлет-преобразование опорной копии сигнала, с искусственно внесённым программным способом доплеровским сдвигом частоты. При этом минимум разностного сигнала будет достигаться тогда и только тогда, когда искусственно вносимый доплеровский сдвиг совпадает с истинным значением доплеровского сдвига. По аналогии с методом оценки времени запаздывания используется многоканальная схема, в каждом канале которой своё значение вносимого доплеровского сдвига.

В результате проведённых исследований были синтезированы новые методы оценки таких параметров радиосигналов, как время запаздывания и доплеровское смещение частоты. Предлагаемые методы являются альтернативой классическим методам, использующих многоканальную структуру из согласованных фильтров, и позволяют получить выигрыш в быстродействии (проведённые тесты показали, что временные затраты при использовании новых методов более чем в 4,5 раза меньше, чем при использовании классических методов). По результатам экспериментов были построены кривые описывающие зависимость относительной погрешности выносимой оценки от отношения «сигнал/шум». При этом в каждом случае проводилось накопление выборки и решение о значении оценки принималось либо по максимуму, либо по среднему, либо по моде полученной выборки. Предполагается проведение дальнейших работ, которые будут направлены на синтез алгоритмов оценивания времени запаздывания и доплеровского смещения с учётом всех заявленных преобразований и последующее оценивание эффективности выносимых оценок.

Список литературы

1. Яковлев А.Н. Введение в вейвлет-преобразования: Учеб. пособие. –Новосибирск: Изд-во НГТУ, 2003. – 104 с.
2. Воробьев В.И., Грибунин В.Г. Теория и практика вейвлет-преобразования. – СПб.: Изд-во ВУС, 1999 – 208 с.
3. William B. Case. Wigner functions and Weyl transforms for pedestrian: Department of Physics, Grinnell College, P.O. Box 805,cGrinnell, Iowa, 50112.
4. Андреева О. М., Маругин А. С., Пыко С. А. Статистическая теория радиотехнических систем: учеб. пособие в 3 ч. Ч. 1 / под общ. ред. проф. В. П. Ипатова и проф. Ю. Д. Ульяницкого. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2015. 146 с.

ИССЛЕДОВАНИЕ СТАБИЛЬНОСТИ ПРИЕМНИКА РАДИОТЕЛЕСКОПА В РЕЖИМАХ ПОЛНОЙ МОЩНОСТИ И МОДУЛЯЦИОННОМ

А. Н. КАМЕЛИ[1], Ю. В. ВЕКШИН [2]

[1]*Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)*

[2]*Институт прикладной астрономии РАН*

Аннотация. В параметризации функций третьего и четвёртого порядка, численном моделировании геометрии трёхмерных поверхностей и тел часто используют математический аппарат алгебры кватернионов. В статье демонстрируются возможности и перспективы применение подобного математического аппарата при моделирования фазовых распределений на элементах антенных решёток.

Ключевые слова: антенные устройства, радиофизика, распространение радиоволн, численное моделирование

Введение

Радиометром называется приемное устройство, предназначенное для измерения мощности шумового радиоизлучения космических источников. В идеале приёмник должен обладать постоянными коэффициентами усиления и стабильной шумовой температурой, усилительные каналы должны иметь линейные характеристики в широком динамическом диапазоне. Нестабильность параметров приёмника приводит к ухудшению реальной точности регистрации радиоастрономических сигналов [1].

Наиболее простой схемой радиометра является радиометр полной мощности, который измеряет суммарную мощность излучения источника и шумов антенны с приемником. Теоретически, он должен обладать наилучшей чувствительностью, но на параметры реального радиометра будут влиять такие параметры, как нестабильности коэффициентов усиления каналов, дрейф нуля усилителя постоянного тока, нестабильность собственных шумов приемной системы. Для устранения этих недостатков используются модуляционные радиометры, в которых применяется периодическая подача на вход приёмника стабильного источника шумового сигнала. Однако это приводит к уменьшению времени измерения источника и ухудшению теоретической чувствительности.

Радиометр состоит из приёмника и системы регистрации. Для исследований стабильности приёмника без влияния системы регистрации необходимо проводить измерения независимым эталонным измерителем мощности. С этой целью в работе реализована регистрация выходных сигналов приёмника с помощью датчика мощности R&S NRP8SN в режимах полной мощности и модуляционном.

Исследуемый приемник и его чувствительность

В работе исследовался приемник X-диапазона длин волн (8.18-9.08 ГГц) для радиотелескопа РТ-32 [2], схема измерений представлена на рис. 1. Приемник состоит из криогенного блока, содержащего охлаждаемые малошумящие усилители (МШУ), блока преобразования частот (БПЧ), блока генераторов шума (БГШ). МШУ необходимы для усиления слабых СВЧ-сигналов при минимальном уровне собственных шумов. Охлаждение усилителей и элементов высокочастотного волноводного тракта производится в криостате. Сигналы на выходе МШУ усиливаются и фильтруются в БПЧ, а также преобразуются к промежуточной полосе частот, в которой работает система регистрации

сигналов. БПЧ содержит смеситель, гетеродин, усилители, фильтры и модулятор, осуществляющий изменение коэффициента усиления канала в модуляционном режиме с пилот-сигналом. БГШ содержит полупроводниковые генераторы шума на лавинно-пролетных диодах и реализует формирование сравнительно слабого сигнала для амплитудной калибровки приемника и мощного сигнала для компенсации собственных шумов антенны и приемника [3]. Сигналы от БГШ вводятся в тракт приемника через направленный ответвитель в криоблоке.

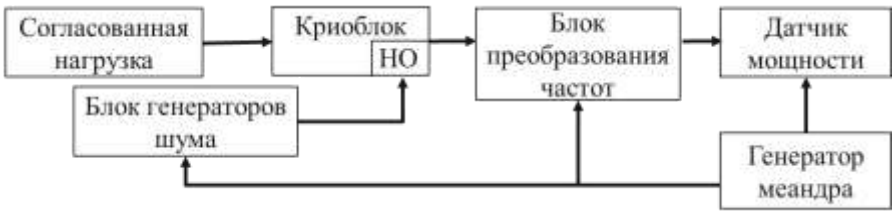


Рис. 1. Схема измерения стабильности приемника в лаборатории

Приемник позволяет работать как в режиме полной мощности, так и в модуляционном режиме с пилот-сигналом. Изменение режима работы приемника осуществляется в программе управления на компьютере. В модуляционном режиме с пилот-сигналом в один полупериод на вход приёмника подается сигнал от антенны, а в другой – сигнал ГШ компенсации с одновременным ослаблением коэффициента усиления в БПЧ на 10 дБ. В режиме полной мощности на вход приёмника все время подается сигнал от антенны, ГШ компенсации выключен. В лабораторных экспериментах на вход приемника устанавливалась согласованная нагрузка с комнатной температурой.

В среде LabView разработана программа, реализующая автоматизированное измерение мощности выходного сигнала приёмника с помощью датчика мощности R&S NRP8SN на задаваемом интервале усреднения τ и запись результатов в файл. Реализовано синхронное детектирование модулированного выходного сигнала приёмника и запись отсчетов в разные массивы в зависимости от уровня сигнала модуляции – TTL меандра с частотой 1 кГц. В датчике мощности R&S NRP8SN использован режим «Trace time» и запуск по сигналу модуляции. Измерения мощности проводились со скоростью 100 кВыб/с, а затем усреднялись отдельно для каждого полупериода на интервале 0.1 с.

Для сравнения двух типов радиометров проводились измерения чувствительности – минимально обнаруживаемого приращения антенной температуры, при котором среднее значения приращения выходного сигнала равно среднеквадратичному отклонению его флуктуаций. Чувствительность ΔT радиометра полной мощности в общем случае равна [1]:

$$\Delta T = (T_A + T_R) \cdot \sqrt{\frac{1}{\Delta F \cdot \tau} + \left(\frac{\Delta G}{G}\right)^2}, \quad (1)$$

где T_A– шумовая температура антенны, T_R – шумовая температура приемника. ΔF – полоса пропускания приемника, а $\Delta G/G$ – относительная нестабильность коэффициента усиления (для идеального случая равна нулю).

Чувствительность модуляционного радиометра с пилот-сигналом рассчитывается по формуле [4]:

$$\Delta T = 2 \cdot \frac{M}{M-1} \cdot \frac{(T_A + T_R)}{\sqrt{\Delta F \cdot \tau}}, \quad (2)$$

где M – коэффициент ослабления модулятора, в данном случае $M = 10$. 2% времени измерения исключалось за счет переходных процессов при переключении сигналов.

Измерения проводились для двух случаев: для криоблока при комнатной температуре $T_ф = 295$ K и охлажденного до температуры $T_ф = 20$ K. При комнатной температуре шумовая температура приемника $T_R = 130$ K, при охлаждении криоблока $T_R = 10$ K. На вход приёмника подключалась согласованная нагрузка с комнатной температурой T_H. Для сравнения стабильности в разных режимах приведем относительное изменение шумовой температуры системы $T_{сист} = (T_H+T_R)$ как среднеквадратичное отклонение выходного сигнала приёмника, делённое на его среднее значение. Расчётные по формулам (1–2) и измеренные значения $\Delta T/T_{сист}$ для полосы приёмника $\Delta F = 900$ МГц при интервале усреднения $\tau = 0.1$ с для различных длительностей измерения приведены в таблице 1.

Таблица 1

Отношение чувствительности к шумовой температуре системы $\Delta T/T_{сист}$

		Расчётное значение	Измеренное значение (время измерения 60 с)	Измеренное значение (время измерения 60 мин)
Криоблок не охлажден	Режим полной мощности	$1{,}05\cdot10^{-4}$	$1{,}17\cdot10^{-4}$	$1{,}96\cdot10^{-4}$
	Модуляционный режим	$2{,}4\cdot10^{-4}$	$2{,}74\cdot10^{-4}$	$2{,}97\cdot10^{-4}$
Охлажденный криоблок	Режим полной мощности	$1{,}05\cdot10^{-4}$	$3{,}55\cdot10^{-4}$	$4{,}01\cdot10^{-4}$
	Модуляционный режим	$2{,}4\cdot10^{-4}$	$2{,}97\cdot10^{-4}$	$3{,}2\cdot10^{-4}$

При «тёплом» криоблоке при времени измерения 60 с относительные чувствительности в режимах полной мощности и модуляционном близки к расчетным, чувствительность в модуляционном режиме с пилот-сигналом в 2.3 раза больше, чем в режиме полной мощности в соответствии с теорией. При «холодном» криоблоке относительная нестабильность в режиме полной мощности больше, чем при «теплом», возможно из-за механических и температурных колебаний при работе микроохладителя криоблока. В модуляционном режиме относительные чувствительности в разных режимах примерно одинаковы за счет уменьшения флуктуаций коэффициента усиления приёмника.

Анализ стабильности сигнала с применением дисперсии Аллана

Среднеквадратичное отклонение временного ряда отсчетов выходного сигнала приемника не позволяет различить тип шума или дрейфа. Шумовые сигналы на выходе могут иметь различную спектральную плотность мощности. Белый шум имеет равномерную спектральную плотность мощности. Фликкер-шум вида $1/f$ – электронный шум, наблюдаемый практически в любых аналоговых электронных устройствах. Но могут присутствовать также флуктуации вида $1/f^{\alpha}$, при которых шум не стационарен. Для таких случаев может потребоваться временной или частотный анализ шумов и дрейфов.

Для рассматриваемого приемника произведен анализ с применением дисперсии Аллана [5], который позволяет определить оптимальное время накопления его выходного сигнала, при котором достигается минимальный уровень дисперсии шума [6]. Также с помощью такого анализа можно исследовать флуктуации вида $1/f^{\alpha}$.

Расчет отклонения Аллана выходного сигнала приемника X-диапазона, нормированного на среднее значение, в зависимости от времени усреднения приведен на рисунке 2. В случае с «теплым» криоблоком при малых значениях интервала усреднения τ

преобладает белый шум (отклонение Аллана уменьшается для обоих режимов). Модуляционный режим с пилот-сигналом имеет в 2.3 раза большее значение отклонения Аллана, что соответствует теории. Минимальный уровень шума для режима полной мощности достигается при времени усреднения $\tau = 3$ с. В модуляционном режиме нестабильность частично компенсируется, белый шум проявляется на более длительных интервалах накопления $\tau = 40$ с, что является оптимальным временем. На интервалах накопления более 5 с чувствительность в модуляционном режиме выше, чем в режиме полной мощности за счет устранения нестабильности коэффициента усиления приемника.

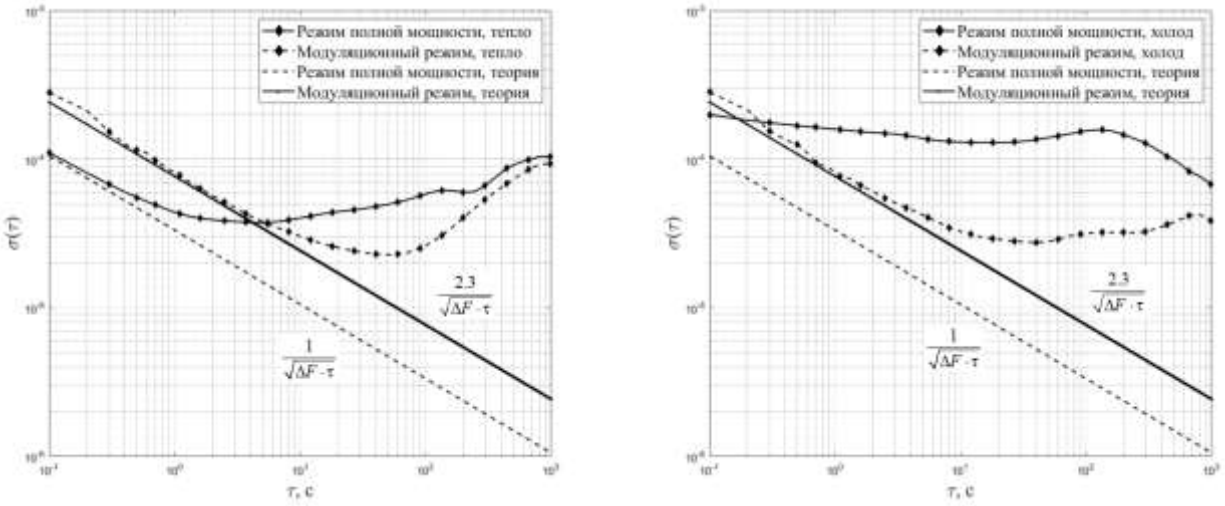


Рис. 2. Относительное отклонение Аллана выходного сигнала приемника в разных режимах

В случае с «холодным» криоблоком в режиме полной мощности относительная нестабильность больше, чем с «теплым» и превышает расчетную, возможно из-за колебаний микроохладителя криоблока. В зарубежных публикациях также отмечается ухудшение стабильности при охлаждении, в том числе из-за свойств полупроводников [7]. Модуляционный режим показывает лучшую стабильность, и его оптимальное время накопления $\tau = 30$ с. Умножив относительную чувствительность на шумовую температуру приемника, получим оценку минимальной чувствительности приемника в градусах K. Минимальное отклонение Аллана для режима полной мощности в «тепле» составляет $4.5 \cdot 10^{-3}$ K, в «холоде» $1.5 \cdot 10^{-3}$ K, для модуляционного режима в «тепле» $3 \cdot 10^{-3}$ K, в «холоде» $2.7 \cdot 10^{-4}$ K. Таким образом, чувствительность приёмника в «холоде» лучше, несмотря на повышенную относительную нестабильность, за счет уменьшения шумовой температуры.

Список литературы

1. Краус Дж. Д. Радиоастрономия. Пер. с английского, под ред. В.В.Железнякова. — М.: Сов радио, 1973 — 456 с.
2. Иванов Д. В. Ипатов А. В. Ипатова И. А. Мардышкин В. В. Михайлов А. Г. Приемники радиоинтерферометрической сети КВАЗАР // Труды ИПА РАН. — СПб.: ИПА РАН, 1997. — Вып. 2. С.247–248.
3. Кольцов Н. Е., Мардышкин В. В., Маршалов Д. А., Евстигнеев А. А. Результаты модернизации приемных каналов S/X диапазонов длин волн на радиотелескопах комплекса «Квазар-КВО» // Труды ИПА РАН. — 2012. — Вып.23. — С. 224–229.
4. Ипатов А. В., Кольцов Н. Е. Радиометры: Учебное пособие. Спб.: Изд-во СПбГЭТУ”ЛЭТИ”, 2007, 92 с.
5. Riley W. J. Handbook of Frequency Stability Analysis // NIST Special Publication 1065. — 2008. — 136 p.

6. Векшин Ю. В., Лавров А. П., Хвостов Е. Ю. Методика и результаты исследований амплитудной и фазовой стабильности супергетеродинного радиоастрономического приемника // Радиотехника. — М.: Радиотехника, 2018. — №1. — С. 24–30.
7. Gallego J. D., Lopez-Fernandez I., Diez C., Barcia A. Experimental results of gain fluctuations and noise in microwave low-noise cryogenic amplifiers // – Noise in Devices and Circuits II / ed. by F. Danneville, F. Bonani, M. J. Deen, M. E. Levinshtein. – Proc. of SPIE. – 2004. – Vol. 5470. – P. 402–413.

ПОМЕХОУСТОЙЧИВОСТЬ ПРИЕМА OTFS-СИГНАЛОВ В УСЛОВИЯХ БЫСТРЫХ ЗАМИРАНИЙ

КОВАЛЕВ Д.М., СЕРГИЕНКО А.Б.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Модуляция в ортогональном частотно-временном пространстве (OTFS) имеет преимущества в сцена-риях с высокой мобильностью из-за разреженного и квазистатичного представления канала. Данные каналы в практических условиях имеют многолучевую структуру с быстрыми замираниями. При ис-следованиях они часто аппроксимируются моделью плоских замираний, что не позволяет учитывать особенности OTFS, связанные с разнесением символов как по частоте, так и по времени. В результате анализа были получены теоретические и смоделированные кривые помехоустойчивости, которые подтверждают устойчивость OTFS к быстрым замираниям, а также демонстрируют выигрыш в диапа-зоне значений максимального доплеровского сдвига, отличного от нуля в широком диапазоне.

Ключевые слова: модуляция OTFS, область «задержка-доплеровский сдвиг», модель канала с быстрыми замираниями.

Введение

В условиях высокой мобильности: высокоскоростные поезда, беспилотный транспорт, самолеты и т.д., а также в беспроводных сетях 6G связь страдает из-за доплеровского разброса на высоких частотах. В таких случаях активно используемое в 4G LTE ортогональное частотное мультиплексирование (OFDM) не является надежным вариантом из-за интерференции между несущими и возрастающей сложности канала связи [1, 2].

В настоящее время активно изучается система модуляции в ортогональном частотно-временном пространстве (Orthogonal Time-Frequency Space, OTFS), которая устойчива к изменяющимся во времени каналам, присущим упомянутым сценариям работы [3].

OTFS использует для передачи данных область задержки-доплеровского сдвига (Delay-Doppler, DD). Преимуществом использования этой области является ее способность преобразовывать канал со случайными замираниями в частотно-временном интервале в квазистационарный канал с разреженным представлением. [1, 2, 3].

При данном типе модуляции в условиях высоких доплеровских разбросов кадр имеет достаточно большую длительность, при которой в реальных сценариях работы характеристики канала перестают иметь сильную статистическую связь. При этом в большинстве работ, посвященных анализу OTFS в каналах с замираниями, эти каналы часто описывают упрощенными импульсными характеристиками, которые не отражают реальных изменений состояния канала на такой длительности кадра.

Целью данной работы является исследование помехоустойчивости приема OTFS-сигналов в каналах с быстрыми замираниями, плохо аппроксимирующимися моделью плоских замираний.

Модель системы

В данной статье рассматривается система OTFS на основе дискретного преобразования Зака (Discrete Zak Transform, DZT) в канале с многолучевыми замираниями и аддитивным белым гауссовым шумом (АБГШ). Схема преобразования показана на рисунке 1.

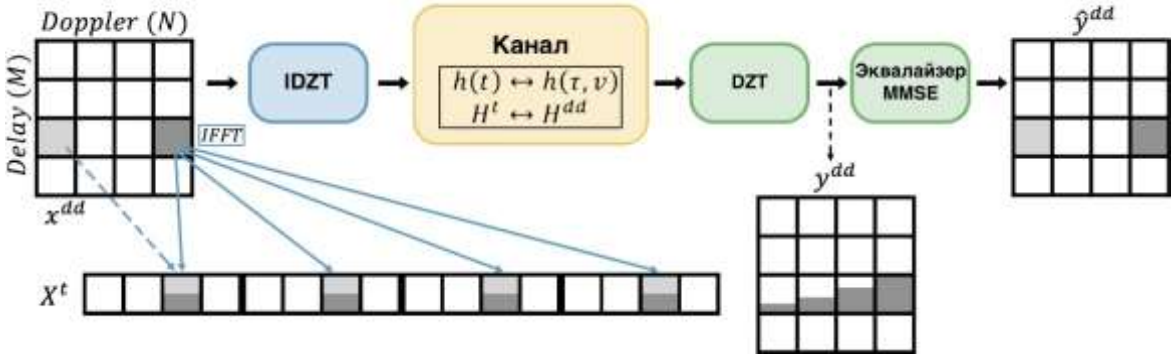


Рис. 1. Структурная схема OTFS-модуляции на основе преобразования Зака

OTFS-кадр в DD-области состоит из MN символов данных на сетке $M \times N$ [4]:

Ось временной задержки: M ячеек $l = 0 \dots M - 1$.

Ось доплеровского сдвига: N ячеек $k = 0 \dots N - 1$.

Символы данных берутся из алфавита $\mathbb{A}$ с n_b битами на символ — квадратурная амплитудная модуляция (КАМ), фазовая манипуляция и т. д.

OTFS-кадр преобразуется из DD-области во временную область с помощью обратного DZT, которое сводится к обратному дискретному преобразованию Фурье вдоль оси доплеровского сдвига и последовательного вывода вдоль оси временной задержки:

$$X^t = (F_N^H \otimes I_M)x^{dd} = V^H x^{dd}, \qquad (1)$$
$$x^{dd} = [x_1, x_2 \dots x_M]^T: \; x_i = \left[x[i,1], x[i,2] \dots \; x[i,N]\right],$$

где x^{dd} – вектор символов данных в DD-области, F_N– матрица быстрого преобразования Фурье размерности N, $I_M \in C^{M \times M}$ – единичная матрица, $\otimes$ – произведение Кронекера, V – матрица DZT-преобразования, определяющая процесс модуляции символа в OTFS.

Сигнал (1) после прохождения через канал связи во временной области имеет вид

$$Y^t = H^t X^t + w = H^t (F_N^H \otimes I_M)x^{dd} + w, \qquad (2)$$

где w – шумовая составляющая (АБГШ), $H \in C^{MN \times MN}$– эффективная матрица канала, H^t во временной области и H^{dd} в DD-области, определяющая свойства канала с замираниями.

Для восстановления сигнала в DD-области к сигналу (2) применяются обратные преобразования, и соотношение «вход-выход» для DD-области имеет следующий вид:

$$y^{dd} = (F_N \otimes I_M)H^t(F_N^H \otimes I_M)x^{dd} + (F_N \otimes I_M)w = H^{dd}x^{dd} + V^H w. \qquad (3)$$

Наличие замираний в канале приводит к необходимости выравнивания полученного сигнала. Искажения, вносимые H^t в (3), могут быть компенсированы эквалайзером:

$$\hat{y}^{dd} = Gy^{dd} = GH^{dd}x^{dd} + GV^{H}w, \tag{4}$$

где G – матрица, определяющая алгоритм линейного выравнивания. Для оценки канала с минимальной среднеквадратической ошибкой (Minimum Mean Square Error, MMSE):

$$G = \left(H^{H}H + \frac{\sigma_w^2}{\sigma_x^2} I_{MN}\right)^{-1} H^{H}, \qquad \hat{G} = \left(\hat{H}^{H}\hat{H} + \frac{\sigma_w^2}{\sigma_x^2} I_{MN}\right)^{-1} \hat{H}^{H}, \tag{5}$$

где σ_w^2 – средняя мощность шума, σ_x^2 – средняя мощность символов данных.

Теоретическая частота битовых ошибок

Для определения теоретической частоты битовых ошибок (Bit Error Rate, BER) необходимо определить отношение сигнал шум (ОСШ) на выходе эквалайзера $\gamma_{m,n}^{out}$, которое затем может быть использовано для пересчета BER на основе вероятности битовой ошибки в АБГШ-канале.

Для КАМ модуляции с алфавитом $\mathbb{A}$ размера K вероятность битовой ошибки в АБГШ-канале может быть достаточно точно аппроксимирована выражением (8.15) из [5]:

$$p_b(\gamma) \cong 4\left(\sqrt{K} - 1\right)/\left(\sqrt{K}\log_2 K\right) \sum_{i=0}^{\sqrt{K}/2-1} Q\left((2i+1)\sqrt{\gamma\, 3\log_2 K/(K-1)}\right), \tag{6}$$

где $Q(\cdot)$ – Q-функция, $\gamma = E_b/N_0$ – ОСШ на бит.

Выражение (6) с усреднением $E\{\cdot\}$ по всем возможным реализациям канала с замираниями и заданными корреляционными свойствами будет иметь следующий вид:

$$BER = E\left\{\frac{1}{MN}\sum_{n=0}^{N-1}\sum_{m}^{M-1} p_b\left(\gamma_{m,n}^{out}\right)\Big|_H\right\}. \tag{7}$$

Для MMSE эквалайзера полезная мощность выровненного сигнала, согласно (4), равна

$$P_{n,m}(x) = |GH|^2\sigma_x^2. \tag{8}$$

Полная мощность на выходе эквалайзера:

$$\begin{aligned} P_{n,m}(y) = E\{y^{dd}(y^{dd})^{H}\} &= GH\; E\{x^{dd}(x^{dd})^{H}\}\, H^{H}G^{H} + GV^{H}\; E\{ww^{H}\}\, VG^{H} \\ &= GHH^{H}G^{H}\sigma_x^2 + GG^{H}\sigma_w^2. \end{aligned} \tag{9}$$

Тогда ОСШ на выходе эквалайзера с учетом (8) и (9):

$$\gamma_{m,n}^{out} = \left[\frac{|GH|^2\sigma_x^2}{GHH^{H}G^{H}\sigma_x^2 + GG^{H}\sigma_w^2 - |GH|^2\sigma_x^2}\right]_{m,n}, \tag{10}$$

где $\gamma_{in} = \sigma_x^2/\sigma_w^2$ – ОСШ на входе эквалайзера.

Поскольку H является эрмитовой матрицей и может быть разложена по собственным значениям $(\lambda_i : i = 0{,}1 \dots MN - 1)$ и векторам (U_i), как $H = U\Lambda U^H$, выражение (10) при наличии идеальной оценки (5) канала преобразуется следующим образом:

$$\gamma_{m,n}^{out} = \gamma_{in}\,[U\ \mathrm{diag}(|\lambda_i|^2)\ U^H]_{m,n} = \gamma_{in}[(F_N \otimes I_M)(H^t)^H\ H^t(F_N^H \otimes I_M)]_{m,n}. \quad (11)$$

Из (11) следует, что выходное значение ОСШ определяется собственными значениями эффективной матрицы канала H, которые характеризуют его состояние. Эта взаимосвязь может помочь понять, почему OTFS дает эффект разнесения в каналах с быстрыми замираниями, в отличие от каналов, аппроксимированных моделью плоских замираний.

OTFS-символ преобразуется из DD-области во временную область, распределяясь по всей длительности кадра (рис. 1), чтобы каждый символ испытывал больше состояний канала, и логично предположить, что OTFS модуляция способна использовать разнообразие свойств канала во временной области для повышения помехоустойчивости при наличии быстрых замираний.

В OTFS при идеальной оценке канала ОСШ одного символа (m, n) рассчитывается с учетом вариаций состояния канала во времени (11), что приводит к снижению чувствительности данного вида модуляции к быстрым замираниям и позволяет получить эффект разнесенного приема.

Результаты моделирования

Для оценки помехоустойчивости было проведено компьютерное моделирование для различных значений максимального доплеровского сдвига f_D в рэлеевских каналах [4] с многолучевыми замираниями. При моделировании использовался кадр $M \times N = 32 \times 32$, модуляция КАМ-4, оценка состояния канала проводилась по встроенному пилот-сигналу с оптимальной мощностью [3], для выравнивания использовался MMSE эквалайзер.

На рисунке 2 показаны смоделированные зависимости BER от ОСШ на бит. Параметром кривой является значение нормированного максимального доплеровского сдвига $f_D N/\Delta f$.

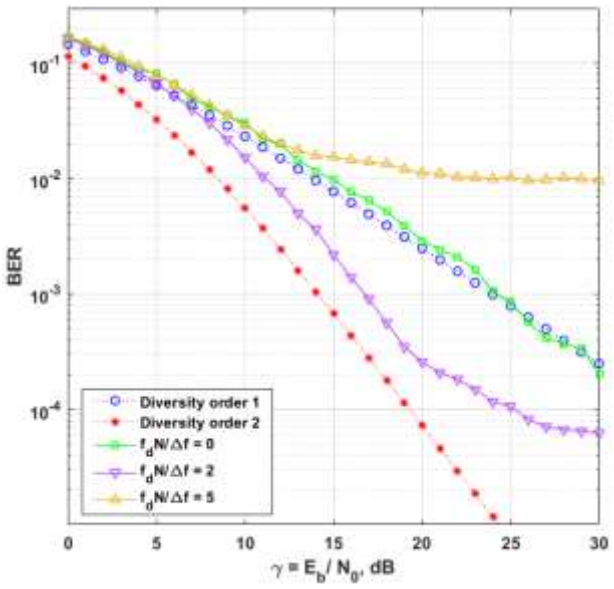


Рис. 2. Помехоустойчивость OTFS в каналах с замираниями

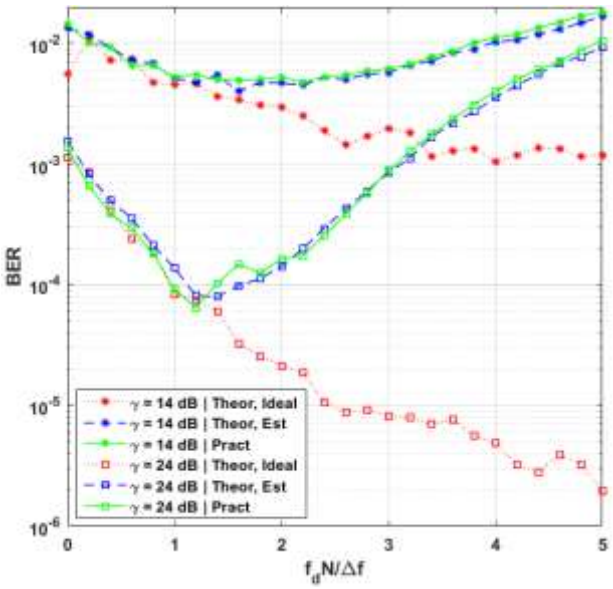


Рис. 3. Зависимость BER от максимального доплеровского сдвига

Увеличение (в некоторых пределах) максимального доплеровского сдвига в рэлеевском канале приводит к уменьшению BER, в том числе относительно канала с плоскими замираниями ($f_D N/\Delta \mathrm{f} = 0$), что отражает возможность OTFS получить выигрыш в помехоустойчивости при вариациях состояния канала на длительности кадра.

На рисунке 3 показаны зависимости BER от максимального доплеровского сдвига при фиксированном ОСШ и теоретические зависимости, построенные по выражениям (7), (10) при идеальной оценке всей эффективной матрицы H^t и при оценке по пилот-сигналу (5). При практической оценке эффект разнесения перестает проявляться при больших допплеровских сдвигах из-за ошибки аппроксимации характеристики канала.

Заключение

В канале с быстрыми замираниями поведение кривых помехоустойчивости OTFS отличается от упрощенных моделей плоских замираний, которые часто используются при анализе. Вариабельность канала во времени при значениях нормированного доплеровского сдвига менее 3 позволяет в некотором диапазоне ОСШ снизить BER на один порядок, как в теоретически аппроксимированных зависимостях, так и в смоделированных. Когда нормированный доплеровский сдвиг становится больше 3, характеристики приема ожидаемо начинают ухудшаться, что свидетельствует о сложности исправления вносимых искажений только за счет структуры OTFS.

Список литературы

1. S. Kumar, S. Deshwal, “A Review on Ortogonal Time Frequecny Space Modulation,” WCNC-2021: Workshop on Computer Networks & Communications Poster: 2021
2. B. Park, H. M. Kwon and H. G. Ryu, “SFFT-Based OTFS Communication System Robust to High Doppler and Long Delay Channel,” *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, Korea (South), 2020, pp. 850-853, doi: 10.1109/ICTC49870.2020.9289247.
3. D. Kovalev, A. Sergienko, “Optimization of Pilot Signal Power in OFDM-based OTFS system,” 27th International Conference on Digital Signal Processing and its Applications (DSPA), Moscow, Russian Federation, 2025.
4. B. Sklar, “Rayleigh fading channels in mobile digital communication systems Part I: Characterization,” IEEE Communications Magazine, 1997, vol. 35, pp. 136-146. doi: 10.1109/35.620535.

5. M.K. Simon, M.S. Alouini, “Digital Communication over Fading Channels: A Unified Approach to Performance Analysis. 1st ed,” New York: John Wiley & Sons, 2000.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АЛГОРИТМОВ РАСПОЗНАВАНИЯ ТИПА ВНУТРИИМПУЛЬСНОЙ МОДУЛЯЦИИ В СРЕДСТВАХ РАДИОМОНИТОРИНГА

МИХАЙЛОВ Д.В.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Прием радиоимпульсов средствами радиотехнического мониторинга ведется в условиях априорной неопределенности относительно типа их модуляции. Для осуществления демодуляции тип модуляции должен быть установлен путем обработки импульса. За последние три десятилетия было разработано множество алгоритмов распознавания типа внутриимпульсной модуляции, которые можно разделить на три большие группы: алгоритмы, на основе метода максимума правдоподобия, алгоритмы на осно-ве признаков типа модуляции и алгоритмы, использующие нейронные сети с глубоким обучением. В настоящем докладе рассмотрен ряд существующих алгоритмов каждой из групп и приведен их сравнительный анализ по критериям вероятности правильного распознавания, числа распознаваемых ти-пов модуляции и сложности.

Ключевые слова: радиомониторинг, распознавание, цифровая модуляция

Распознавание типа модуляции можно представить, как задачу различения сигналов. Оптимальным по критерию минимизации ошибки подходом к задаче различения в предположении равной вероятности присутствия каждого из сигналов является метод максимума правдоподобия (МП) [1], заключающийся в вычислении отношения правдоподобия (ОП) для каждой из гипотез и выбора в качестве решения той гипотезы, для которой значение ОП максимально.

На практике применение метода МП усложняется наличием априорно неизвестных параметров сигнала – начальной фазы, частоты и амплитуды несущего колебания, а также последовательности передаваемых символов. Тогда при различении сигналов, выполняется либо усреднение ОП по плотностям вероятности (ПВ) параметров, рассматриваемых как случайные величины (СВ), либо предварительное оценивание параметров по методу МП с дальнейшей подстановкой полученных оценок в выражение для ОП [2]. Оба этих подхода, а также полученные на их основе квазиоптимальные подходы применялись к задаче распознавания типа модуляции [2 - 4]. В [2] был также предложен комбинированный подход, названный авторами гибридным тестом отношения правдоподобия (ГТОП) и предполагающий получение оценки для амплитуды, частоты и фазы и усреднение ОП по символам рассматриваемого типа модуляции.

Один из алгоритмов, основанных на ГТОП, был предложен в [5] в 2017 году. Данный алгоритм предполагает прием сигнала несколькими датчиками с целью повышения точности распознавания. Для последовательности символов, принятых каждым датчиком, требуется оценить комплексный канальный коэффициент h, учитывающий амплитуду и фазу несущего колебания и дисперсию шума σ^2. Предложено использовать квазиоптимальный метод получения оценок, называемый онлайн алгоритмом Каппѐ. Суть метода состоит в формировании начальной оценки по первому принятому символу сигнала и ее последовательном уточнении после приема каждого следующего символа. По

сравнению с обработкой всей принятой последовательности целиком, данный подход позволяет не только ускорить процесс распознавания, за счет получения оценки параметров в ходе приема сигнала, но и, как показано в [5], снизить вычислительную сложность. Затем с использованием полученных оценок вычисляются ОП и модуляция, соответствующая наибольшему из них, принимается в качестве решения.

Вторая группа алгоритмов использует различные признаки, характерные для разных типов модуляции для их распознавания. В их число входят кумулянты последовательности символов [6], сигнальные созвездия [7], статистические характеристики амплитуды, частоты и фазы [8] и т.д.

Кумулянтами случайного процесса (СП) называются коэффициенты разложения логарифма его характеристической функции в ряд Маклорена. Один из алгоритмов, основанных на использовании кумулянтов был предложен в работе [9] в 2020 году. В его основе лежит сравнение десятичных логарифмов кумулянтов с заранее установленными верхним и нижним порогами. Каждому типу модуляции соответствуют свои значения порогов и потому в какой диапазоне лежат величины логарифмических кумулянтов вид модуляции может быть распознан.

На первом шаге рассчитываются кумулянты принятой последовательности символов, затем вычисляются их логарифмы. На основании значений логарифмов сначала определяется является ли используемая модуляции фазовой манипуляцией (ФМ) или квадратурной амплитудной модуляцией (КАМ), после чего уточняется порядок модуляции.

Другим алгоритмом, основанным на признаках, является предложенный в [10] метод распознавания типа модуляции сигнала на фоне шума, распределение которого является не гауссовым, а шире – принадлежит семейству устойчивых распределений. В основе алгоритма лежит тот факт, что радиосигнал, смешанный с шумом, является циклостационарным СП, то есть таким СП, математическое ожидание и автокорреляционная функция которого являются периодическими. Для циклостационарных процессов может быть вычислена функция, известная в англоязычной литературе как cyclic correntropy function (CCF), которая используется как признак типа модуляции. Для распознавания рассчитывается CCF принятой реализации и вычисляется ее коэффициент корреляции (КК) с эталонной CCF каждого из предполагаемых типов. В качестве решения принимается тип модуляции, для которого значение КК оказалось наибольшим.

В последние годы работы, посвященные распознаванию типа модуляции с помощью метода МП или заранее определенных признаков, активно вытесняются статьями, предлагающими подходы к распознаванию на основе нейронных сетей (НС) с глубоким обучением (ГО). В подобных НС признаки типов модуляции, которые в дальнейшем будут использованы для распознавания, формируются на этапе обучения в скрытых нейронных слоях. В качестве входных данных для формирования признаков может требоваться только принятая реализация [11] или принятая реализация, дополненную полученными на ее основании признаками, такими как сигнальные созвездия [12] или кумулянты [13].

В качестве примера будет рассмотрена НС, использующая в качестве исходных данных отсчеты принятой реализации и их циклические кумулянты [14], архитектура которой предложена в [13]. Алгоритм ее работы выглядит следующим образом. Принятая последовательность символов поступает в слой извлечения признаков, извлекающим основные признаки сигнала. Выход данного слоя подается на восемь параллельно функционирующих капсул, каждая из которых состоит из двух сверточных слоев.

Выходные данные капсул поступают в полностью связанный слой (ПСС), в котором они преобразуются в вектор, размерностью $1\times L$, где L – число распознаваемых типов модуляции. Данный вектор подается на оконечный слой softmax, выносящий решение о типе модуляции.

Основные показатели качества рассмотренных алгоритмов сведены в таблицу 1

Таблица 1

Основа алгоритма	Вероятность правильного распознавания	Отношение сигнал/шум, дБ	Число накапливаемых символов
метод МП	0,99	2	250
логарифмы кумулянтов	0,99	5	500
CCF	0,99	5	
глубокое обучение	0,9	2	32 768

Среди представленных алгоритмов подходы, основанные на методе МП и на CCF обладают наибольшей универсальностью и могут быть использованы для распознавания, вообще говоря, любых видов цифровой модуляции. В случае первого метода для расширения перечня распознаваемых типов потребуется только вычислить большее количество ОП в процессе распознавания, а в случае второго типа – увеличить число заранее рассчитанных эталонных CCF, при этом выражения для расчета ОП и CCF остаются неизменными.

При использовании НС, для изменения перечня распознаваемых типов потребуется повторное обучение, а для его расширения – увеличение размерности ПСС, то есть, изменение архитектуры. В случае алгоритма на основе логарифмических кумулянтов потребуется разработка новых порогов или даже новых кумулянтов, так как значения используемых кумулянтов могут оказаться слишком близки друг к другу для новведенных типов, что приведет к их частому перепутыванию.

Как видно из таблицы 1, алгоритм на основе МП демонстрирует высокую точность 0,99 при накоплении всего 250 символов, тест алгоритма выполнялся для типов модуляции КАМ-8, КАМ-16, ФМ-8 и ФМ-16. Однако, как показано в [5], он обладает высокой сложностью, пропорциональной $N\cdot K\cdot\sum_{l=1}^{L}M_l$, где N – число накопленных символов, K – число датчиков, L – число распознаваемых типов, M_l - порядок l-го типа модуляции. Таким образом, даже при накоплении. В связи с этим, обработка даже сравнительно небольшого числа символов может потребовать существенных временны́х затрат.

Выборочные оценки кумулянтов представляют собой линейные комбинации выборочных оценок моментов разных порядков, например выражение оценки кумулянта $c_{2,2}$ имеет вид:

$$\hat{c}_{2,2}=\mathrm{cum}\left\{y(n)^*,y(n)^*\right\}=\mathrm{M}\left\{\left(y(n)^*\right)^2\right\}-\left(\mathrm{M}\left\{y^2(n)\right\}\right)^2-2\left(\mathrm{M}\left\{y(n)\cdot y(n)^*\right\}\right)^2,$$

где $\mathrm{cum}\{\cdot\}$ - операция вычисления кумулянта, $\mathrm{M}\{\cdot\}$ - операция усреднения (оценки математического ожидания, "*" - операция комплексного сопряжения. Таким образом алгоритм обладает невысокой сложностью, пропорциональной числу символов.

Как видно из таблицы 1 алгоритм может обеспечить хорошую точность распознавания сигналов ФМ-{2, 4, 8} и КАМ-{8, 16, 32, 64} при ОСШ более 5 дБ. Увеличение числа накапливаемых символов до 10 000 позволяет в дополнение к перечисленным типам распознавать также КАМ-128 и КАМ-256, достигая средней точности 0,99 при ОСШ –3 дБ.

Выборочная оценка CCF вычисляется в соответствии с выражением [10]:

$$V(\omega,\tau) = 1/N \sum_{n=0}^{N} 1/\left(\sigma\sqrt{2\pi}\right)\cdot\exp\left(-\left(x(n)-x(n+\tau)\right)^2/\left(2\sigma^2\right)\right)\cdot\exp\left(-j2\pi\omega n\right),$$

где N – число принятых символов, n – номер символа ω – циклическая частота, τ – некоторая задержка, j – мнимая единица. Алгоритм может обеспечить точность распознавания типов модуляции ФМ-2, ФМ-4, КАМ-8, КАМ-16 и КАМ-32 0,99 при ОСШ, равном 5 дБ. Однако в [10] авторы не сообщают число принятых символов, обеспечивающих данную точность. Это не позволяет полностью корректно сравнить данный алгоритм с остальными, однако он представляет интерес как один из немногих алгоритмов, рассчитанных на негауссов шум.

Использование глубокого обучения позволяет распознавать достаточно широкий перечень типов модуляции, в который входят ФМ-{2, 4, 8}, КАМ-{16, 64, 256}, ФМ-4 со смещением и модуляцию с минимальным частотным разнесением. Однако, как видно из таблицы 1, вероятность правильного распознавания невелика по сравнению с другими алгоритмами даже при накоплении большого числа отсчетов.

По итогам сравнения рассмотренных алгоритмов в качестве наиболее предпочтительного, при работе в условиях гауссова шума, предлагается отметить метод, основанный на логарифмах кумулянтов. Ввиду своей низкой вычислительной сложности и способности обеспечивать хорошую вероятность правильного распознавания накопив сравнительно небольшое количество символов, он позволяет быстро выносить верное решение, не используя при этом сложных систем обработки сигналов.

Список литературы

1. Перов А.И. Статистическая теория радиотехнических систем. СПб.: Радиоэлектроника, 2003. 399 с.
2. Prokopios Panagiotou, Achilleas Anastasopoulos, Andreas Polydoros. Likelihood ratio tests for modulation classification, DOI: 10.1109/ACSSC.1995.540877
3. Keith M. Chugg, Chu-Sieng Long, Andreas Polydoros. Combined Likelihood Power Estimation and Multiple Hypothesis Modulation Classifcation, DOI: 10.1109/MILCOM.2000.904013
4. Fahed Hameed, Octavia A. Dobre, Dimitrie C. Popescu. On the Likelihood-Based Approach to Modulation Classification, DOI: 10.1109/TWC.2009.12.080883
5. Berkan Dulek. Online Hybrid Likelihood based Modulation Classification using Multiple Sensors, DOI: 10.1109/TWC.2017.2704124
6. M.R. Mirarab, M.A. Sobhani. Robust Modulation Classification for PSK /QAM/ASK Using Higher-order Cumulants, DOI: 10.1109/ICICS.2007.4449591
7. Jacques P. Mouton, Melvin Ferreira, Albertus S.J. Helberg. A comparison of clustering algorithms for automatic modulation classification. Expert Systems with Applications, 2020, vol. 151, DOI: 10.1016/j.eswa.2020.113317
8. E.E. Azzouz, AK. Nandi. Automatic identification of digital modulation types, DOI: 10.1016/0165-1684(95)00099-2
9. Ahmed K.Ali, Ergun Erçelebi. Algorithm for automatic recognition of PSK and QAM with unique classifier based on features and threshold levels, DOI: 10.1016/j.isatra.2020.03.002
10. Tales V. R. O. Câmara, Arthur D. L. lima1, Bruno M. M. Lima, Aluisio I. R. Fontes, Allan de M. Martins, Luiz F. Q. Silveira. Automatic Modulation Classification Architectures Based on Cyclostationary Features in Impulsive Environments, DOI: 10.1109/ACCESS.2019.2943300
11. Muhammad Zeeshan Mumtaz, Muhammad Khurram, Muhammad Adnan ,Adnan Fazil. Autonomous Modulation Classification using Single Inception Module based Convolutional Neural Network, DOI: 10.1109/IBCAST51254.2021.9393285

12. Hui Han, Zhijian Yi, Zhigang Zhu, Lin Li, Shuaige Gong, Bin Li, Mingjie Wang. Automatic Modulation Recognition Based on Deep-Learning Features Fusion of Signal and Constellation Diagram, DOI: 10.3390/electronics12030552
13. John A. Snoap, Dimitrie C. Popescu, James A. Latshaw, Chad M. Spooner. Deep-Learning-Based Classification of Digitally Modulated Signals Using Capsule Networks and Cyclic Cumulants, DOI: 10.3390/s23125735
14. Chad M. Spooner. Classification of Co-Channel Communication Signals Using Cyclic Cumulants, DOI: 10.1109/ACSSC.1995.540605

ОБЗОР СПОСОБОВ ПОСТРОЕНИЯ РЛС С СДЦ

Хмыров И.К.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. В данной работе описаны основные принципы радиолокации и влияние на процесс радиолокационного наблюдения пассивных помех. Для борьбы с ними применяется селекция движущихся целей на основе эффекта Доплера. Приводятся краткие характеристики РЛС разных типов, применительно к решаемым задачам.

Ключевые слова: селекция движущихся целей, радиолокационные системы, эффект Доплера.

В основе радиолокации лежит получение информации о объектах наблюдения (положение в пространстве, скорость движения и др.) с помощью радиоволн. Процесс обнаружения объектов, измерение их координат и параметров движения называют радиолокационным наблюдением. Оборудование, использующееся для этого, называют радиолокационными станциями (РЛС). Основной принцип работы РЛС заключается в излучении зондирующего сигнала, часть энергии которого достигает объекта наблюдения, называемого целью. Энергия, перехваченная целью, отражается во многих направлениях, включая направление на радиолокатор. Отраженный сигнал принимается антенной РЛС, что позволяет принять решение о наличии или отсутствии цели. В первом случае могут быть определены координаты цели и другая информация о ней. Радиолокационными целями могут быть как самолеты, суда или ракеты, так и люди, птицы, атмосферные осадки и космические объекты [1].

Однако, мешающие объекты, не являющиеся целями, препятствуют получению достоверной информации о цели. Мешающими объектами могут выступать: подстилающая поверхность, возвышенности, гидрометеоры, здания. Отраженные радиосигналы от таких объектов называют пассивными помехами. Воздействие пассивных помех может превышать уровень собственных шумов приемника, в некоторых случаях и полезный сигнал цели, что сильно затрудняет его радиолокационное наблюдение. Методы борьбы с помехами основаны на различии характеристик сигналов, отраженных целью и мешающими отражателями. Обычно эти различия связаны с неодинаковой протяженностью и положением их в пространстве, также скоростью движения. Для улучшения соотношения сигнала к помехе на входе приёмника ключевым фактором является повышение пространственной избирательности радиолокационной станции. Это достигается за счёт:

Увеличения разрешающей способности – необходимо, чтобы размер разрешаемого элемента (объёма или площади) максимально соответствовал физическим габаритам цели. Сужения диаграммы направленности антенны (ДНА) – это уменьшает попадание помех из

боковых лепестков и повышает концентрацию энергии в направлении цели. Расширения спектра зондирующего сигнала – использование широкополосных или сложных сигналов (например, ЛЧМ) улучшает разрешение по дальности и снижает влияние помех.

Таким образом, оптимизация условий наблюдения цели требует комплексного подхода: ужесточения пространственной селективности (за счёт узкой ДНА) и улучшения разрешающей способности. Это позволяет минимизировать влияние помех и повысить эффективность обнаружения [1,2,3].

При проектировании радиолокационной станции ключевой задачей является оптимальный подбор параметров зондирующего сигнала и характеристик согласованного фильтра в приемнике. Такая оптимизация направлена на достижение максимальной вероятности обнаружения цели при заданных требованиях к вероятности ложной тревоги и заданном соотношении сигнала цели к сумме помехи и шума приемника. Один из наиболее эффективных способов обработки сигналов на фоне пассивной помехи является введение в систему, помимо СФ, второго фильтра, подавляющего частотные составляющие спектра помехи, которые отличаются от сигнальных на величину разности их доплеровских смещений. Такая фильтрация, называется селекцией движущихся целей (СДЦ). Основные задачи СДЦ включают обнаружение самолетов на фоне отражений от земных объектов и облаков различного происхождения для наземных РЛС; выявление целей на низкой высоте на фоне отражений от поверхности Земли для бортовых РЛС; и определение наземных транспортных средств на фоне отражений от стационарных объектов и дорог для автомобильных РЛС [2].

Доплеровские методы СДЦ основаны на различии доплеровских смещений частоты выделяемого полезного сигнала цели и пассивных помех, обусловленном различием радиальных скоростей цели и мешающих объектов. Для простоты расчётов предполагается, что помехообразующие объекты неподвижны. В этом случае доплеровский сдвиг частоты F_v возникает только за счёт радиальной скорости цели относительно этих неподвижных объектов.

$$F_v = \frac{2v_r}{\lambda_u}$$

где $\lambda_и$ – длина волны излучаемого РЛС сигнала [2].

Чтобы выделить доплеровское смещение, нужно сравнить частоту излученного и отраженного сигналов, если цель имеет ненулевую радиальную скорость относительно излучателя, частота отраженного от нее сигнала будет отличаться от частоты излучаемого РЛС сигнала на величину доплеровского смещения. Следовательно, в случае, если источники пассивных помех неподвижны относительно РЛС, отраженный от них сигнал будет иметь нулевое смещение [2].

Данный принцип проще всего использовать в РЛС непрерывного излучения, т.к. отправленный и принятый сигналы в таких системах существуют одновременно. Это позволяет сравнивать частоты отправленного и полученного сигналов. Однако, несмотря на высокую точность измерения скорости, он не применяется на практике в радиолокационных системах с обнаружением движущихся целей из-за строгих требований к разделению приёмной и передающей антенн и отсутствия дифференциации по расстоянию. Поэтому в основном применяются импульсно-доплеровские РЛС. Известно, что импульсные РЛС с функцией выделения сигналов движущихся объектов различаются в зависимости от частоты повторения зондирующих импульсов. РЛС с частотами ниже 50

кГц относят к низкочастотным, в то время как станции с частотой выше 50 кГц относят к импульсно-доплеровским системам. Есть также РЛС со средней частотой повторения, лежащпоказаний ни по скорости, ни по дальности, однако использование разных частот может улучшить эти параметры [1].

Разделяют импульсно-доплеровские системы с внешней и внутренней когерентностью. В случае внутренней когерентности, в системе есть внутренний генератор опорных колебаний, с частотой которых сравнивается частота полученного сигнала. Это помогает определить величину доплеровского смещения, и следовательно скорость цели. Внутренне когерентные РЛС делят на истинно-когерентные и псевдокогерентные. Истинно-когерентные РЛС излучают когерентные последовательности импульсов, являющиеся отрезками одного и того же высокочастотного сигнала, что обеспечивает одинаковую начальную фазу всех импульсов. Источником такого сигнала часто является кварцевый генератор. Частота опорного сигнала непосредственно сравнивается с частотой, полученного сигнала, но сравнение происходит на промежуточной частоте, что достигается с помощью умножения частоты опорного генератора [2,3].

В псевдокогерентных РЛС источник опорных колебаний – это когерентный гетеродин, фазируемый колебаниями генератора передатчика в течение длительности импульса. Когерентный гетеродин работает на частоте сравнения, на которой происходит выделение доплеровского сдвига. Обычно это промежуточная частота приемника. Системы называют псевдокогерентными из-за того, что начальная фаза колебаний хаотично меняется, что приводит к необходимости осуществления фазирования когерентного гетеродина.

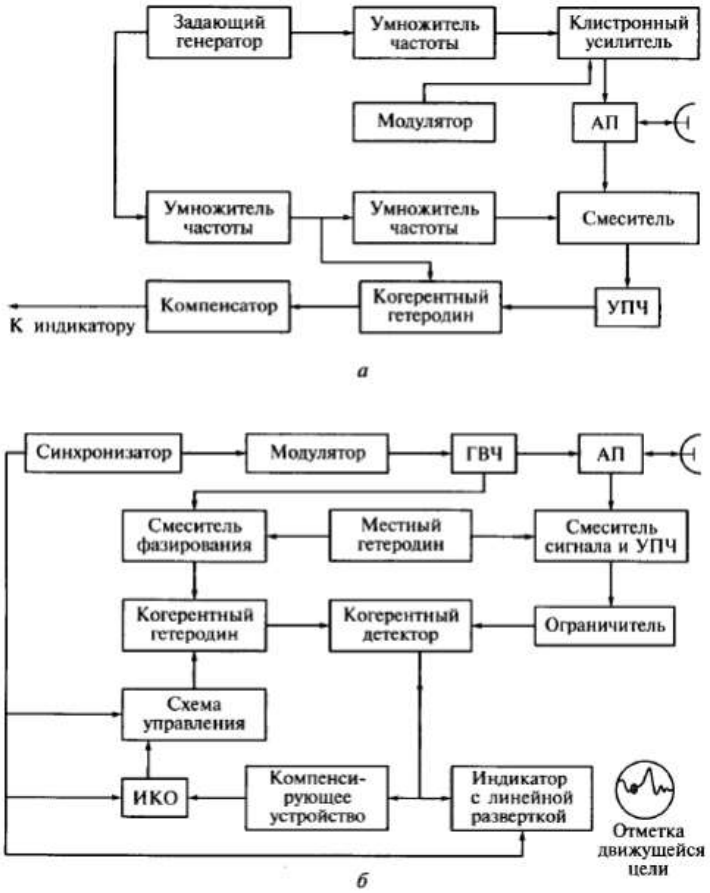


Рис.1. Структурные схемы импульсно-доплеровских РЛС с внутренней когерентностью

Если при фазировании когерентного гетеродина разность фаз когерентного гетеродина и фазирующих колебаний сохраняется от импульса к импульсу постоянной, то амплитуда импульсов сигнала от неподвижных объектов будет сохраняться постоянной, что обеспечивает эффективное подавление таких сигналов при последующей фильтрации в компенсирующем устройстве СДЦ [2].

В системах с внешней когерентностью частота сигнала, пришедшего на вход приемного устройства сравниваются с частотой сигнала, отраженного от неподвижных целей, которые находятся в том же сегменте разрешения, что и предполагаемая цель. Данные системы более актуальны для использования на борту движущихся объектов (например, корабли) ввиду того, что здесь нет необходимости подстраивать частоту под собственную скорость РЛС. Однако, в связи с тем, что сигнал, отраженный от неподвижных объектов, меняется во времени, эффективность данных РЛС немного ниже, чем РЛС с внутренней когерентностью. В результате использование в качестве опорного сигнала, отраженного от неподвижных объектов, на экране индикатора оператор видит лишь изображение движущихся объектов. Данный способ также предполагает возможность использования компенсирующего устройства, что увеличивает эффективность системы. Существенным недостатком данных систем является то, что при отсутствии фона, отсутствует и опорный сигнал, при этом СДЦ не происходит. Избавиться от этого недостатка можно с помощью отключения СДЦ [2].

Список литературы

1. Справочник по радиолокации / Под ред. М.И. Сколника. Пер. с англ. под общей ред. В.С. Вербы. В 2 книгах. Книга 1 Москва: Техносфера, 2014. - 672 с., ISBN 978-5-94836-381-3.
2. Радиотехнические системы / Ю.М. Казаринов и др.; под ред. Ю.М. Казаринова. М.: Издательский центр «Академия», 2008. 529 с.
3. Финкельштейн М.И. Основы радиолокации: учебник. 2-е издание. –М.: Радио и связь,1983. – 536 с.

ГИПЕРСПЕКТРАЛЬНАЯ СИСТЕМА ДЛЯ ДИАГНОСТИКИ ПОВРЕЖДЕНИЙ ТКАНЕЙ

ШАТКОВСКАЯ М.С.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Гиперспектральное сканирование (HSI) представляет собой перспективную технологию получения изображений, содержащих спектральную информацию для каждого пикселя сцены. В диагностике повреждённых тканей, включая ожоги и обморожения, HSI обеспечивает высокую детализацию, позволяя визуализировать структуру и состав материала без контакта с объектом. Несмотря на значительный потенциал, широкое применение гиперспектральных систем в прикладной медицине сдерживается рядом технических ограничений, включая громоздкость оборудования, длительное время съёмки, сложность интерпретации и высокую стоимость. В настоящей статье рассматриваются существующие методы формирования гиперспектральных изображений, анализируются принципы работы одного из наиболее известных решений — системы TIVITA 2.0, а также обсуждаются возможные пути технической оптимизации гиперспектральных устройств для повышения их пригодности в прикладных задачах визуального контроля состояния тканей.

Ключевые слова: гиперспектральное сканирование, медицинская визуализация, спектр.

Современные 1. Введение

Современные методы визуализации, такие как макросъёмка, термография и допплерография, не обеспечивают достаточного спектрального охвата для комплексной оценки состояния повреждённых тканей. Гиперспектральная визуализация позволяет получать спектрально-пространственные данные, где каждому пикселю соответствует отражательный или пропускной спектр, что особенно важно при диагностике неоднородных, термически или ишемически изменённых участков. Ключевые физиологические параметры определяются в следующих диапазонах: оксигенация — 570–590 нм и 740–780 нм; содержание гемоглобина — 530–590 нм и 785–825 нм; перфузия — 655–735 нм и 825–925 нм; содержание воды — 880–900 нм и 955–980 нм. Несмотря на высокую информативность, широкое внедрение HSI ограничено громоздкостью оборудования, сложностью оптики, длительным временем съёмки и чувствительностью к условиям освещения. Это подчёркивает необходимость инженерной адаптации систем под реальные условия применения.

2. Технические методы формирования гиперспектральных изображений

Гиперспектральное изображение представляет собой трёхмерный массив данных, называемый гиперспектральным кубом, где двум пространственным координатам (x, y) соответствует множество спектральных каналов (λ). Методы получения таких данных различаются по принципу разложения спектра и организации съёмки. Ниже рассмотрены основные подходы к формированию HSI-изображений с технической точки зрения.

2.1 Сканирующие методы формирования гиперспектральных изображений

Сканирующие методы формируют гиперспектральный куб последовательно — по строкам, длинам волн или освещению.

Push-broom-системы используют щелевую апертуру и одновременно регистрируют все спектры только для одной строки сцены. Для построения полного куба необходим контроль перемещения объекта или камеры, что ограничивает применение стабильными условиями — например, в лабораториях или на конвейерах.

Фильтровое сканирование реализуется с помощью сменных механических или перестраиваемых фильтров (AOTF, LCTF), размещённых перед или после объектива. Съёмка происходит поочерёдно для каждой длины волны, к тому же спектральная характеристика источника должна быть равномерной. Метод обеспечивает точный спектральный отбор, но требует времени на перестройку.

Сканирование с перестраиваемым источником достигается за счёт поочерёдного включения LED или изменения частоты излучения твердотельного источника. При этом детектор остаётся неподвижным, а фильтры не используются. Метод отличается простотой оптики, но требует точной синхронизации и равномерности освещения.

2.2 Snapshot-системы

Snapshot-системы формируют гиперспектральный куб за один кадр, что делает их особенно подходящими для съёмки динамичных сцен. Это достигается за счёт различных инженерных решений: например, использования микрофильтров на сенсоре, где каждый пиксель или группа пикселей регистрирует сигнал в определённой длине волны, либо с помощью призменных или дифракционных элементов, оптически разделяющих спектры разных участков изображения. Также применяются мультиапертурные схемы, позволяющие одновременно регистрировать несколько спектральных компонент. Хотя такие системы обеспечивают высокую скорость съёмки, они ограничены по спектральному

разрешению, требуют точной калибровки и сложной постобработки. Тем не менее, они находят применение в портативных решениях, где приоритетами являются компактность и оперативность.

Таблица 1

Сравнение методов формирования гиперспектральных изображений

Метод	Преимущества	Недостатки	Применение
Push-broom (линейное сканирование)	Высокое спектральное разрешение (до 2–5 нм);	Требует точного позиционирования; Чувствителен к движениям объекта;	Используется преимущественно в лабораторных условиях.
Сканирование через фильтры (Filter wheel, AOTF, LCTF)	Более предсказуемый спектральный профиль; Возможность использования мощного стационарного освещения.	Время сканирования ограничивается из-за переключения фильтров; Потери света на фильтрах; Чувствительность к движению объекта.	Системы доминируют в коммерческих медицинских приборах благодаря компактности
Сканирование с помощью перестраиваемого источника излучения (LED-матрицы, лазеры, твердотельные источники)	Простая архитектура при отсутствии фильтров перед объективом; Высокая отношение сигнал/шум за счёт прямого освещения; Возможность компактной реализации.	Чувствительность к движению; Необходима точная синхронизация между источником и сенсором; У каждого канала своя пространственная неоднородность освещения	Подход активно используется в прикладных и инженерных прототипах, особенно там, где важна модульность системы.
Snapshot-системы	Мгновенный захват спектральной информации (важно при движении объекта). Высокая скорость формирования кадра	Значительные искажения при пространственно-спектральной деконволюции; Требуется трудоёмкая предобработка	Перспективны для портативных решений, но пока уступают по точности сканирующим устройствам.

2.3 Особенности формирования гиперспектрального куба

При формировании гиперспектрального изображения важно учитывать несколько технически значимых факторов. Неустойчивое освещение может вызвать искажения спектров, а длительная съёмка повышает риск смещения объекта. Оптические системы подвержены хроматическим аберрациям, возникающим из-за различий в преломлении света на разных длинах волн. Кроме того, большой объём данных (сотни мегабайт на один куб) требует продуманной системы хранения и передачи. Выбор метода съёмки всегда связан с необходимостью баланса между разрешением, временем регистрации и условиями применения.

3. Пример реализованной HSI-системы: TIVITA 2.0

TIVITA® 2.0 (Diaspective Vision GmbH, Германия) — стационарная гиперспектральная система для визуализации поверхностных тканей, работающая в диапазоне 500–1000 нм. Она включает широкоспектральный осветитель, перестраиваемый интерференционный фильтр и камеру с разрешением 640×480 пикселей. Устройство формирует гиперспектральный куб из 100 спектральных полос с шагом около 5 нм. Область съёмки

составляет до 20×15 см, размер пикселя — около 0,3–0,5 мм². Полный цикл регистрации длится около 6,4 секунд, при этом каждое изображение записывается отдельно, что требует неподвижности объекта. Объём одного куба достигает 120 МБ. Система чувствительна к движениям и нестабильному освещению, а также отличается крупными габаритами (~20 кг), высокой стоимостью и ограниченным полем зрения.

При всей технологической проработанности TIVITA остаётся примером высокоточной, но стационарной и специализированной установки, слабо приспособленной для использования вне контролируемых условий. Вопросы мобильности, упрощения конструкции и снижения стоимости становятся особенно актуальными в контексте расширения спектра практических применений гиперспектральной визуализации.

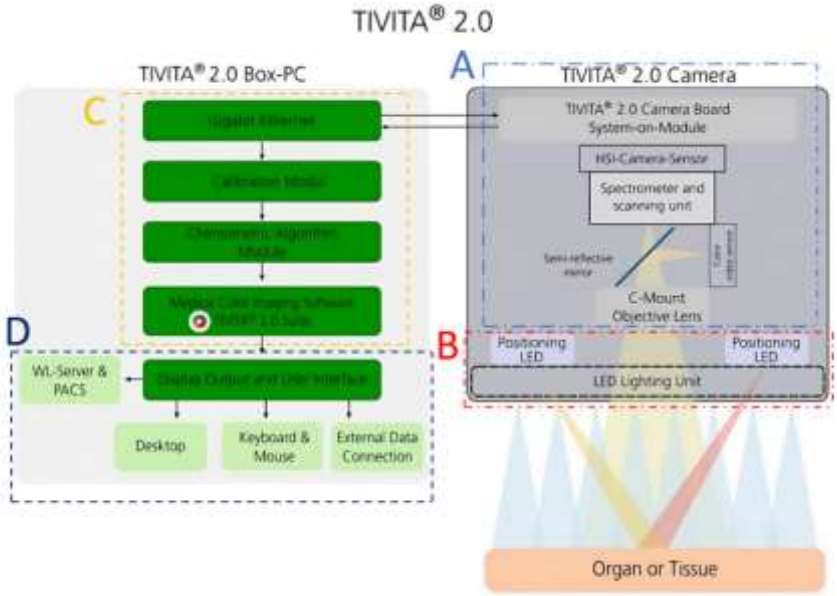


Рисунок 1. Структурная схема системы TIVITA 2.0: A — гиперспектральная камера и цветной видеосенсор (сенсор, спектрометр, перестраиваемый фильтр, объектив, зеркало, цветовой видео сенсор); B — система позиционирования и модуль освещения (направляющие светодиоды, LED-подсветка); C — блок обработки данных (модули калибровки, спектрального анализа, ПО); D — пользовательский интерфейс и внешние соединения (монитор, клавиатура, сеть).

Это определяет направление дальнейших инженерных усилий, направленных на создание более доступных и гибких решений, о чём пойдёт речь в следующем разделе.

4. Пути оптимизации гиперспектральных систем для визуализации повреждённых тканей

Разработка прикладных гиперспектральных систем требует баланса между качеством спектральной информации и инженерной реализацией. Один из путей оптимизации — переход от громоздких фильтрационных систем к компактным источникам с перестраиваемым спектром, например, на базе LED-матриц или AOTF-лазеров. Это упрощает конструкцию, повышает надёжность и снижает стоимость.

Миниатюризация также достигается за счёт современных CMOS-сенсоров и короткофокусной оптики, что позволяет создавать компактные модули, пригодные для использования в полевых и ограниченных условиях.

Полноспектральные кубы из 100–200 каналов избыточны, поскольку для расчёта ключевых параметров (оксигенация, перфузия, водосодержание и др.) используются только

отдельные диапазоны длин волн. Это позволяет сократить число каналов до 20–50 без потери информативности, снизив объём данных и ускорив съёмку.

Наконец, для практического применения системы должны быть устойчивы к внешним воздействиям. Регулируемое освещение, жёсткий корпус, быстрая фиксация и масса менее 5 кг обеспечивают мобильность без ущерба для качества данных и делает возможной работу в нестабильных условиях.

5. Заключение

Гиперспектральная визуализация позволяет получать пространственно-спектральную информацию о состоянии тканей и является перспективным методом для диагностики повреждений с выраженной неоднородностью структуры и оптических свойств.

В работе рассмотрены основные методы формирования гиперспектральных изображений, включая сканирующие и snapshot-системы, а также проанализирована техническая реализация системы TIVITA®. Выявленные инженерные ограничения таких устройств указывают на необходимость оптимизации.

Предложены пути технического совершенствования, включая замену спектральных фильтров на перестраиваемые источники, сокращение числа каналов, миниатюризацию оптики и упрощение интерфейса. Это позволит создавать более доступные и прикладные системы без существенной потери качества данных.

Дальнейшее развитие HSI-технологий требует фокусировки не только на точности, но и на адаптации под условия реального применения, что открывает путь к их интеграции в задачи визуального контроля повреждённых тканей.

Список литературы

1. Lu G., Fei B. Medical hyperspectral imaging: a review // *Journal of Biomedical Optics*. – 2014. – Vol. 19(1). – P. 010901. – DOI: 10.1117/1.JBO.19.1.010901
2. Zhang Y., Wu X., He L., Meng C., Du S., Bao J., Zheng Y. Applications of hyperspectral imaging in the detection and diagnosis of solid tumors // *Transl Cancer Res 2020;9(2):1265-1277 - DOI: http://dx.doi.org/10.21037/tcr.2019.12.53*
3. Bhargava A., Sachdeva A., Sharma K., Alsharif M.H., Uthansakul P., Uthansakul M. Hyperspectral imaging and its applications: A review // *Heliyon*. – Volume 10, Issue 12, 30 June 2024 – DOI: https://doi.org/10.1016/j.heliyon.2024.e33208
4. TIVITA® 2.0 Datasheet. Diaspective Vision GmbH. – 2022. – URL: https://diaspective-vision.com/wp-content/uploads/2020/11/TI-05-MD-03_A_TIVITA_2.0_Datenblatt.pdf (дата обращения: [16.04.2025]).

ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА МОДУЛЬНОГО ДРОНА С ОТКРЫТЫМ ПО

МАНОЕНКО Г.С.

Кузбасский государственный технический университет имени Т.Ф. Горбачева

Аннотация. В статье реализован концепт модульности БПЛА, были подобраны необходимые компоненты, произведены расчеты стоимости компонентов для сборки, созданы 3D – модели самого устройства, а также крепления модулей.

Ключевые слова: модульность, крепление, компоненты, БПЛА, прототип.

Развитие Беспилотные летательные аппараты (БПЛА) активно применяются в геодезии, картографии, мониторинге, логистике и развлекательной сфере (например, гонки дронов). Однако большинство существующих моделей имеют фиксированную конструкцию, что ограничивает их адаптивность к различным условиям.

Модульность дрона позволяет:

- оптимизировать аэродинамику под разные задачи, путем замены роторов (скорость, манёвренность, устойчивость);
- уменьшать габариты для транспортировки и работы в ограниченном пространстве;
- повышать энергоэффективность за счёт замены двигателей;
- оптимизировать устройство под разные задачи путем добавления и/или замены модулей.

Цель проекта.

Разработка и создание прототипа модульного БПЛА, способного адаптироваться под различные условия эксплуатации.

Задачи.

- Анализ существующих решений – изучение конструкций существующих типов корпусов и способов реализации модульности.
- Выбор концепции модульности.
- Подбор компонентов – двигатели, драйверы, датчики и т.д.
- Создание 3д-модели устройства.

В ходе работы было проанализировано несколько способов реализации модульности устройства, таких как подсоединение модулей при помощи разъемов USB Type – C, Type – B, micro USB, крепление с помощью магнитного адаптера, а также крепление типа «Горячий башмак». Было принято решение остановиться на последнем.

Рисунок 1 «Горячий башмак»

Крепление состоит из:

- металлической направляющей (башмака) – имеет Т-образную форму с контактами;
- соответствующего разъема на подключаемом устройстве;
- центрального контакта;
- дополнительных контактов;
- фиксатора (замка) – предотвращает случайное отсоединение.
- электрического контакта;
- центрального контакта, замыкающего цепь, позволяя контролеру отправлять сигнал на модуль.

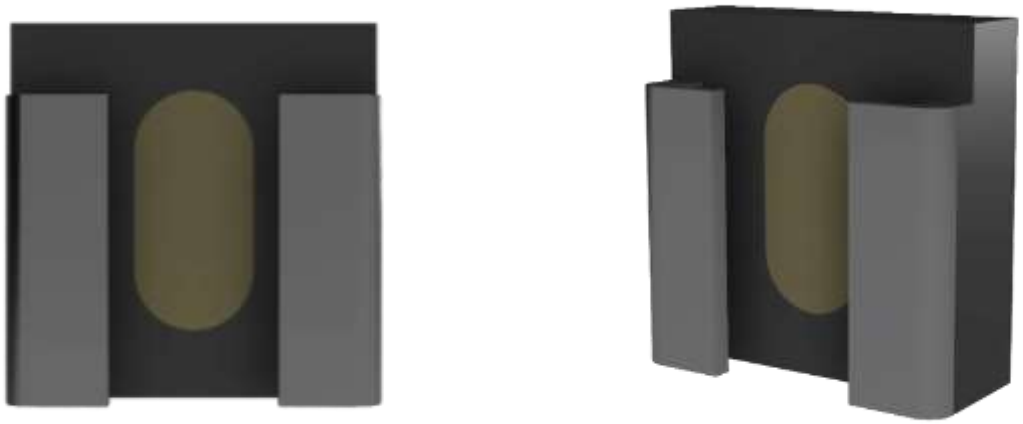


Рисунок 2 3D – модель крепления

Рисунок 3 «Горячий башмак» дрона

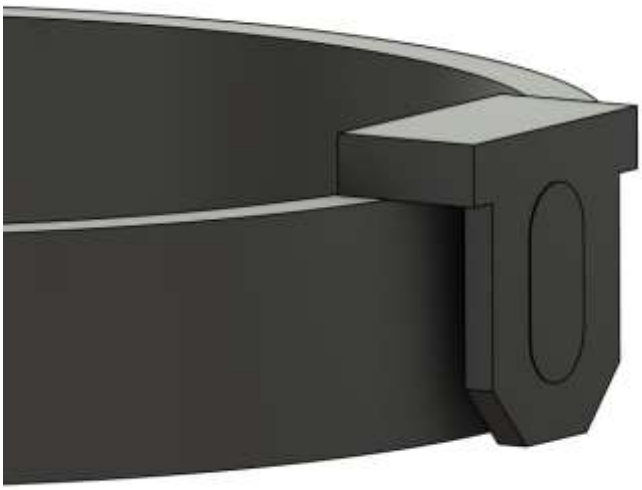


Рисунок 4 Крепление модуля дрона

В качестве полетного контроллера была выбрана плата Arduino Nano, для дальнейшей разработки собственно ПО

Драйвера (ESC) и все комплектующие были подобраны в соответствии с характеристиками подранного мотора.

Так же для реализации автопилота была выбран модуль гироскопа и акселерометра MPU6050.

При рассмотрении форм – факторов различных дронов первоначально выбор пал на трикоптер из-за его небольших габаритов и меньшей стоимости по сравнению с иными вариантами. Но необходимость установки дополнительного сервопривода в «хвосте» заставило изменить решение в связи с дополнительным усложнением конструкции и уменьшением универсальности. Потому окончательным вариантом стал более востребованный и всем известный квадрокоптер.

3D – модель, созданная при помощи Autodesk Fusion.

Рисунок 5 3D - модель прототипа

Датчики, которые планируется в дальнейшем установить/возможно будет установить на дрона:

- манометр
- гигрометр
- анемометр
- датчик освещенности.

В ходе работы был проведен анализ, подбор и расчет стоимости компонентов для сборки прототипа. Результаты представлены в таблице.

Таблица 1

Расчет стоимости компонентов

Наименование	Кол-во	Цена
Драйвер(ESC)	8	3360
Бесщеточный двигатель	8	4568
Литий-полимерный аккумулятор 2S(3800 мАч)	1	1000
ПЛК(Arduino Nano)	1	350
Винты	12	1350
Рама	1	0
Приемник ELRS 2.4G	1	1000
Модуль гироскопа MPU6050	1	300
		11928

Заключение.

Проанализировав ключевые моменты проектирования и сборки устройства, в результате были произведены расчеты, подобраны необходимые компоненты, созданы 3D - модели, так же реализован концепт модульности и спроектирован новый тип крепления для установки расширений.

Список литературы

1. Мультироторные дроны и БПЛА: все, что нужно знать // rccopter.ru URL: https://rccopter.ru/blogs/knowledge/multirotornye-drony-i-bpla-vse-chto-nuzhno-znat?srsltid=AfmBOoryqKtzpswZaUpLzGXGckwcnja5qVYTWRNoYpYlx-vfZFUJ4ofm (дата обращения: 30.03.2025).
2. https://habr.com/ru/articles/227425/(дата обращения: 30.03.2025)
3. https://habr.com/ru/articles/197878//(дата обращения: 30.03.2025)

СЕКЦИЯ ЛИНГВИСТИКА

ОСОБЕННОСТИ БРИТАНСКОГО ЮМОРА

АЛЕКСАНДРОВ К.С., ВАСИЛЬЕВА О.А., БАЛАЕВА К.А.

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация. Статья посвящена британскому юмору как культурному явлению и его формам в современном обществе. Рассматриваются ключевые черты юмора, такие как ирония, самоирония, сарказм, абсурд и юмор неловкости, с примерами из британских телешоу и сериалов. Особое внимание уделено роли юмора в профессиональной среде и контексте искусственного интеллекта. Статья подчеркивает важность культурного контекста для эффективного использования юмора в бизнесе и технологиях.

Ключевые слова: британский юмор, ирония, сарказм, самоирония, культурные особенности, коммуникация.

Юмор — важная часть повседневной жизни британцев. Люди, выросшие в других странах и проживающие вне Великобритании часто неправильно понимают британский юмор. Иногда трудно понять, шутят ли британцы на самом деле. Некоторые поражаются их способности находить смешное практически в любой ситуации, даже в самой печальной. Многие британцы считают, что их чувство юмора уникально, оно более тонкое и развито в большей степени, чем у других народов. Популярный британский драматург Оскар Уайльд выразил эту мысль достаточно ясно, сказав: «It is clear that humour is far superior to humor». В данной цитате обыгрывает различное написание слова «humour» в британском и американском английском, где «humor» — американское написание.

Однако, пожалуй, самое непонятное в британском юморе то, что в нем нет кнопки «выключить». Почти в каждом разговоре между британцами обязательно присутствует та или иная форма иронии, сарказма, подшучивания, недосказанности, самоуничижения, поддразнивания или насмешки. Когда в каждом слове, которым обмениваются британцы, чувствуется скрытый юмор, становится трудно понять, шутит британец или говорит серьезно. Это еще более проблематично, учитывая, что шутки почти всегда передаются с невозмутимым выражением лица. Таким образом, негласное правило гласит: если кто-то с невозмутимым видом говорит что-то абсолютно бессмысленное, он, скорее всего, шутит.

Одной из ключевых черт британского юмора является ирония, часто переходящая в самоиронию. В культуре, где не принято выставлять себя напоказ, умение посмеяться над собой рассматривается как признак зрелости и интеллектуального такта. Самоирония служит своеобразным социальным буфером: она обезоруживает и позволяет сохранить достоинство даже в самых неловких ситуациях. В свою очередь, ирония в более широком смысле — это форма дистанцированного отношения к действительности, способ выразить мнение или критику без грубости и прямоты.

Несмотря на репутацию утончённости, британский юмор не чужд и более приземлённым формам, таким как туалетные и сексуально окрашенные шутки. Особенно распространены намёки (innuendo), построенные на двусмысленности и недосказанности.

Такие шутки подаются с привычной сухостью и серьёзностью, что усиливает их комический эффект. Данный юмор служит способом разрядки, разрушения табу и демонстрации раскованности в рамках социального кода. В популярном в Британии сериале «Офис» звучала фраза «as the actress said to the bishop» [1]. Это выражение «как

актриса сказала епископу» получила основную популярность благодаря персонажу Рики Джервейса Дэвид Брент, который начал использовать ее в популярном британском ситкоме «Офис», а персонаж Стива Кэррелла Майкл Скотт адаптировал ее как «сказала она» в американской версии шоу. Из-за того, что священнослужители проводили много времени с актрисами, пытаясь заставить их очиститься и изменить свой образ жизни, родилась фраза «сказала актриса епископу», намекающая на то, что актрисы признаются в своих регулярных грехах перед священнослужителями.

Сарказм, как более острая форма иронии, также занимает важное место в британском юморе [2]. Его особенность заключается в сдержанности: зачастую он подаётся без яркой интонации или мимики, из-за чего его можно принять за серьёзное высказывание. Такой приём требует от собеседника внимания к контексту и интонационным нюансам. Британский сарказм — это не агрессия, а интеллектуальное высказывание с долей ироничной насмешки, зачастую направленной не только на других, но и на самого говорящего. Примером сарказма может служить следующий отрывок из сериала «Шерлок»:

- *Андерсон:* «Rache» по-немецки означает «Месть». Возможно, она пытается нам что-то сказать.
- *Шерлок:* Да, спасибо за вашу проницательность.

В данном случае Шерлок высказывает свой саркастический комментарий, чтобы проиллюстрировать противоположность своему буквальному высказыванию, т.е. он хочет донести свое раздражение до Андерсона, потому что ясно, что слово «Rache» используется для обозначения чего-то, и очевидно, что Шерлок умен и уже знает это.

Важной чертой британского юмора является так называемая «сухость подачи» (deadpan delivery). Это приём, при котором даже самые абсурдные или остроумные реплики произносятся абсолютно серьёзным тоном. Отсутствие эмоциональной окраски создает комический эффект засчёт контраста между содержанием и формой. Подобный стиль не только подчёркивает интеллектуальность юмора, но и требует от зрителя активного участия — нужно уловить подтекст, чтобы понять, где заканчивается серьёзность и начинается шутка. В британской передаче «Cunk on Earth» ведущая Филомена Канк произносит данную фразу: «Почти невероятно, что до того, как Чарльз Дарвин изобрел теорию эволюции в 1859 году, никто и никогда не эволюционировал. Без него никто из нас не существовал бы сегодня, разве что в виде окаменелостей или гиббонов». Все шоу построено на абсурде и безэмоциональной подаче неуместных реплик и комментариев ведущей к ученым, профессорам и докторам. Шутка основана на абсурде — будто эволюция началась только после Дарвина, что, конечно, нелепо. Сухость подачи усиливает комизм — ведущая говорит это серьёзно, как факт, что создаёт контраст между тоном и бессмыслицей сказанного. Такой приём заставляет зрителя распознать юмор самостоятельно, подчёркивая интеллектуальность и сатирический стиль шоу.

Абсурд занимает в британском юморе особое место. Он проявляется в построении нелепых, нелогичных ситуаций, которые при этом подаются как абсолютно нормальные. За счёт этого достигается эффект разрушения привычных представлений о реальности и здравом смысле. Британский абсурд — это не просто попытка вызвать смех, а способ показать странность повседневных норм, бюрократии, социальных ролей и даже самой идеи порядка [3]. Этот приём делает юмор глубоко философским, хотя на поверхности он может выглядеть просто нелепо. На абсурде строится множество шуток и культуры

британцев, например франшиза фильмов от Монти Пайтона. В фильме «Монти Пайтон и Священный Грааль» произошел такой диалог:

- *Король болотного замка:* Вы убили всего лишь отца невесты.
- *Сэр Ланселот:* Ну, я не хотел.
- *Король болотного замка:* Не хотел? Вы пронзили своим мечом его голову.
- *Сэр Ланселот:* О боже... с ним все в порядке?

Ланселот, убив человека, искренне спрашивает, всё ли с ним в порядке. Комизм возникает из нелогичности реакции и серьёзной подачи, типичной для британского юмора — серьёзное отношение к очевидному абсурду.

Британский юмор также широко опирается на комизм ситуаций и неловкости [4]. Часто источник смеха — не шутка как таковая, а сама обстановка: недопонимание, социальное напряжение, стремление сохранить вежливость в заведомо абсурдных условиях. Такой юмор близок к повседневной жизни, он высвечивает противоречия между правилами поведения и реальными человеческими реакциями. Особенно ценится способность находить смешное в самых обыденных, порой даже болезненных моментах.

- *Роза:* А как он выглядит?
- *Доктор:* Как передатчик, большой и круглый, находится где-то в центре Лондона! Большой металлический предмет круглой формы — колесо или тарелка и он близко отсюда, наверное он невидим…

Доктор встает спиной к London Eye, Роза пристально смотрит за спину доктора

- *Доктор:* Что? *Роза смотрит за спину Доктора, он оборачивается к колесу*
- *Доктор:* Что? *Доктор оборачивается к колесу еще раз*
- *Доктор:* Да что такое? *Роза подмигивает, Доктор оборачивается последний раз*
- *Доктор:* Ааа, прекрасно!

Смех вызывает недопонимание между героями и напряжённое ожидание момента «озарения». Комизм усиливается повторяющимися реакциями и неловкостью, пока Доктор наконец не догадывается. Это типичный британский приём — находить юмор в повседневной ситуации, социальной неуклюжести и запоздалом осознании очевидного.

Наконец, нельзя не отметить интеллектуальность британского юмора. Он насыщен культурными отсылками — к истории, литературе, политике, философии. Такие шутки требуют от зрителя подготовки, они не всегда понятны с первого взгляда, но именно в этом и заключается их ценность. Юмор становится способом поддержания культурной идентичности, маркером принадлежности к определённому кругу, где ценится не столько реакция, сколько способность интерпретировать и понимать сложные смыслы.

Новичкам поначалу может быть сложно понять британский юмор. Если сарказм и ирония не являются частью юмора в вашей родной культуре, их может быть трудно обнаружить, а также они могут быть неверно истолкованы. Культурные отсылки также являются важной частью британского юмора, поэтому для его полного понимания необходим определенный объем базовых знаний.

Помимо этого, юмор может использоваться в работе, переговорах, сфере обслуживания. Юмор играет решающую роль в деловом общении, являясь инструментом установления взаимопонимания, снятия напряженности и создания позитивной рабочей атмосферы. Британский юмор, в частности, характеризуется сухим остроумием и недосказанностью, что может быть эффективным в преодолении стрессовых ситуаций.

В контексте искусственного интеллекта юмор может использоваться для гуманизации технологических взаимодействий, делая их более доступными и менее пугающими.

Появление искусственного интеллекта привело к эволюции делового языка, появились новые термины и жаргон. Филологи могут понаблюдать за тем, как британский юмор адаптируется к этим изменениям, часто высмеивая сложность, а иногда и абсурдность технического языка. Например, использование юмора в учебных материалах или внутренних коммуникациях может сделать внедрение новых технологий искусственного интеллекта более приемлемым для сотрудников. Британский юмор глубоко укоренен в культурных нюансах, которые могут быть как сильной стороной, так и сложной задачей в условиях глобализации бизнеса. Изучая данный вопрос, ученые могут обнаружить, как эти культурные особенности влияют на восприятие и интеграцию искусственного интеллекта. Например, использование юмора в обслуживании клиентов, основанном на ИИ, может улучшить пользовательский опыт, но оно должно быть тщательно адаптировано, чтобы избежать культурного недопонимания. Искусственный интеллект преобразует рабочие места, начиная с автоматизации рутинных задач и заканчивая предоставлением расширенной аналитики. Британский юмор может помочь сотрудникам справиться с неопределенностью и беспокойством, связанными с этими изменениями. Используя юмор в программах обучения ИИ и корпоративных коммуникациях, компании могут создать более устойчивую и адаптирующуюся рабочую силу. Чтобы проиллюстрировать эти моменты, можно проанализировать конкретные примеры, в которых британский юмор эффективно использовался в бизнес-контекстах, управляемых искусственным интеллектом. Например, использование юмористических чат-ботов с искусственным интеллектом в сфере обслуживания клиентов или интеграция юмора в обучающие модули на базе искусственного интеллекта могут дать ценную информацию о практическом применении этого подхода.

В заключение отметим, что британский юмор является не только частью культуры и двигателем искусства, но и предлагает богатый и детальный способ разобраться в сложностях интеграции ИИ в современную бизнес-среду. Используя уникальные качества британского юмора, компании могут разработать более ориентированный на человека и привлекательный подход к внедрению ИИ. Это не только повышает моральный дух сотрудников и удовлетворенность клиентов, но и способствует развитию культуры инноваций и адаптивности. Поскольку искусственный интеллект продолжает развиваться, роль юмора в формировании его интеграции будет оставаться интересной областью изучения как для филологов, так и для профессионалов бизнеса.

Список литературы

1. Where the Saying 'That's What She Said' Actually Came From // culture trip URL: https://theculturetrip.com/europe/united-kingdom/england/articles/where-the-saying-thats-what-she-said-actually-came-from (дата обращения: 05.05.2025).
2. British humour: What exactly is it and how does it work? // The South American URL: https://www.thesouthafrican.com/lifestyle/move-to/south-africans-abroad/british-humour-what-exactly-is-it-and-how-does-it-work/ (дата обращения: 05.05.2025).
3. Understanding British humour – knowing when to laugh and when not to! // Study International URL: https://studyinternational.com/news/understanding-british-humour/ (дата обращения: 05.05.2025).
4. What Makes British Humour Unique // Remitly URL: https://blog.remitly.com/en-gb/culture/british-humour/ (дата обращения: 05.05.2025).
5. Юмор в бизнес-коммуникации: как применять в выступлениях и на переговорах // BusinessSpeach URL:

https://business-speech.ru/blog/yumor-v-biznes-kommunikatsii-kak-primenyat-v-vystupleniyakh-i-na-peregovorakh/?ysclid=mabl7pu75e520136373 (дата обращения: 05.05.2025).

МОДЕЛИ ОБРАЗОВАНИЯ СЛЕНГИЗМОВ В АМЕРИКАНСКОМ КИНОТЕКСТЕ: ЛИНГВИСТИЧЕСКИЙ И СОЦИОКУЛЬТУРНЫЙ АСПЕКТ

КУЩЕНКО А.П.[1], МАЛЫШЕВА В.Н.[1]

[1] *Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)*

Аннотация. В статье рассматриваются основные модели образования сленгизмов на материале американских фильмов 1995–2020 гг. Проанализированы способы словообразования (фонетическое искажение, сложение, редупликация, усечение, конверсия и метафоризация) с точки зрения их лингвистических характеристик и социокультурной функции. Исследование демонстрирует, как сленг в кинотексте отражает идентичность молодежной аудитории, придавая речи персонажей экспрессию, аутентичность и коммуникативную динамику.

Ключевые слова: молодежный сленг, словообразование, кинотекст, редупликация, метафоризация, конверсия, экспрессия.

Введение

Молодежный сленг — динамичная форма лексики, тесно связанная с социокультурными изменениями. В нём находят отражение такие изменения, как трансформация ценностных ориентиров, влияние цифровых технологий и глобализация культуры. Эти факторы способствуют активному заимствованию, переосмыслению и созданию новых лексических единиц в молодёжной среде [10; с. 1].

Американские фильмы играют ключевую роль в его распространении, делая кинотекст репрезентативным источником для изучения механизмов формирования сленговых слов и выражений. Особое влияние сленга проявилось с развитием голливудского кинематографа и глобального медиапространства молодёжи. Согласно данным Motion Picture Association, в 2023 году американская киноиндустрия экспортировала продукции на сумму $22,6 млрд. Кроме того, американские фильмы распространяются более чем в 130 странах, что свидетельствует о высокой степени их востребованности и потенциала воздействия на международную аудиторию, в том числе на молодежь [11].

Тема молодёжного сленга получила широкое освещение в лингвистике, особенно в последние десятилетия (Тамбовцева, 2015; Кулькарни, 2017; Песина, 2024), что связано с его динамической природой и активной ролью в неформальной коммуникации. Д. К. Каминский подробно исследовал структуру и функции сленга в работе «Сленг: структура и функции» (1999). А. Д. Шмелёв также затрагивал вопросы использования сниженной и разговорной лексики, в контексте социолингвистических исследований. В англоязычной традиции классическими считаются труды Э. Партриджа [14,15], в которых сленг рассматривается как важный и самобытный пласт лексики. Существенный вклад также внесли Дж. Грин [16], Т. Торн с работами по современному молодёжному сленгу [17], а также Д. Кристал, осветивший особенности неформальной и разговорной речи [13].

Цель настоящего исследования — выявить продуктивные модели образования сленгизмов и определить их функцию в молодежном дискурсе на примере киноязыка.

Методы и материал исследования

В исследовании использованы методы сплошной выборки, лингвистического описания и структурно-семантического анализа. Материалом стали субтитры и диалоги из американских фильмов: Clueless (1995), Mean Girls (2004) и Bad Boys for Life (2020), охватывающие молодежную лексику за период с 1995 по 2020 годы. Статус принадлежности слова к сленгу устанавливался по словарю Oxford English Dictionary [12]. В общей сложности было отобрано 51 сленговая единица, которые классифицировались по способу образования и уровню языка. Основное внимание уделено практическому функционированию этих выражений в диалогах, их прагматической нагрузке и стилистической роли в кинотексте. Основное внимание уделено практическому функционированию этих выражений в диалогах: анализируется их прагматическая нагрузка, то есть коммуникативная цель использования, а также их стилистическая роль в создании живого, аутентичного киноязыка, приближенного к повседневной разговорной речи.

Фонетический уровень

Сленговые выражения, созданные на этом уровне, характеризуются фонетическим искажением нормативных форм с целью усиления эмоционального эффекта и принадлежности к определённой социолингвистической группе. Среди наиболее распространённых видов фонетических искажений в английском сленге выделяются: замена гласной фонемы, при которой происходит замена одного гласного звука другим с целью стилизации (*byotch* от *bitch*); редукция, при которой происходит ослабление или опущение неударных слогов, часто приводящее к слиянию слов: *gonna* вместо *going to*, *lemme* вместо *let me*; элизия, или выпадение звука или группы звуков (’*cause* вместо *because*, *fam* вместо *family*); консонантная замена — замена одного согласного на другой, обычно с целью подражания произношению субкультур (*dat* вместо *that*).

Примерами образования сленгизмов на фонетическом уровне могут быть следующие:

– *Byotch* (от *bitch*) «стерва» (Mean Girls) — мягкая ироничная альтернатива грубому выражению. Героиня употребляет в фильме данное слово для демонстрации принадлежности к определённой молодежной группе, а также для создания образа дерзкой, но модной девушки.

– *Yuck* «гадость» употребляется в фильме Clueless и является примером слова, образованного путем звукоподражания (ономатопеи), поскольку имитирует спонтанный звук отвращения, который человек издает, когда видит что-то неприятное. *Yuck* представляет собой простое, краткое и эмоциональное междометие. Использование данного слова персонажем показывает, что он не сдерживает эмоций, является честным и импульсивным.

– *Jeez* «черт побери» – используется в фильме Bad Boys for Life и представляет собой пример усечения. Оно произошло от *Jesus* «Иисус» и употребляется как междометие, выражающее эмоциональную реакцию. В английском языке подобные эвфемистические формы замещают религиозно окрашенные слова с целью смягчить высказывание.

Морфемный уровень

Морфемные способы образования сленга включают следующие процессы:

– словосложение, при котором происходит объединение двух и более слов в одно. Образованные слова могут писаться слитно или через дефис. Например, *nutbox* «психушка» [9] образовано сложением *nut* «сумасшедший человек» и *box* «коробка»;

– редупликация, при которой происходит удвоение слога или корня слова с целью придания экспрессивности: *blah-blah*, *goody-goody*, *chi-chi* [8; 3];

– аббревиация, или сокращение слов или словосочетаний: *FWB* ‘***f**riends **w**ith **b**enefits’* [8] «друзья с преимуществами»;

– усечение — удаление части основы слова, обычно конца или начала: *aight* ‘**a**l**right**’ «ладно, хорошо», *sus* ‘**sus**picious’ «подозрительный» [8];

– аффиксация — добавление приставок или суффиксов: *nigger**ish*** «человек, который ведет себя как типичный темнокожий человек»*, jeep**er**s creep**er**s!* «Боже мой!» [8].

Примерами сленговых выражений, образованных на морфемном уровне, являются:

– *airhead* (сложение *air* «воздух» + *head* «голова») «тупица» — метафорически описывает легкомысленного человека (Clueless). Фраза “*Am I an airhead*?”, использованная героиней фильма, представляет собой самокритичный вопрос. *Airhead* часто используется в американском сленге для обозначения легкомысленных девушек. Фраза подчёркивает неуверенность девушки и попытку понять, как ее воспринимают другие. Интонация может быть обеспокоенной, но ироничной.

– *Big jug* (Mean Girls) – выражение, образованное путем сложения основ прилагательного *big* «большой» и существительного *jug* «кувшин». В результате возникает метафорический образ, который используется для описания прибавки в весе с оттенком пренебрежения или насмешки.

– *Dickhead* представляет собой сложное слово, образованное путём объединения двух самостоятельных лексем — *dick* (вульгаризм, обозначающий глупца или грубого человека) и *head* (элемент, обозначающий тип личности). Подобная структура типична для разговорного и уничижительного сленга (*air**head**, meat**head**, bone**head**)*. В фильме Bad Boys for Life это выражение используется в конфликтной ситуации. Оно демонстрирует грубость, резкость и отсутствие фильтров в общении между персонажами.

Лексический уровень

Для описания процессов словообразования, происходящих на лексическом уровне, можно опереться на классификации семантических изменений, предложенные Л. Блумфилдом и Д. Бланком. Блумфилд выделяет такие типы изменений, как расширение, сужение, метафора, метонимия, деградация и возвышение значения. В отличие от него, Бланк разрабатывает когнитивно-семантический подход: перенос по сходству (метафора), по смежности (метонимия), логический перенос, эвфемизация и др.

– *Fetch* (в значении «стильно») — пример семантического сдвига (метонимии) (Mean Girls). Изначальное значение *fetch* «принести», «добыть» смещается к значению «модный», «привлекательный» на основе идеи, что то, за чем «стоит пойти», является желанным. Употребление слова служит демонстрацией стремления персонажа диктовать модные тренды, выделиться среди сверстников.

– *I’m toast* (Clueless) – фраза образована путем метафоризации и используется в значении «мне конец». *I’m toast* — эмоциональная фраза, выражающая страх провала или наказания. Слово *toast* приобретает переносное значение: как подгоревший тост — всё безнадежно. Фраза демонстрирует свойственную подросткам склонность к драматизации.

– Фраза *I dusted your ass* представляет собой пример сленгового выражения, образованного посредством метафоризации. Глагол *to dust* в своем прямом значении означает «вытирать пыль». Однако в разговорной американской речи это слово приобретает новое, метафорическое значение — «победить, разгромить кого-либо». Такая семантическая переоценка происходит по ассоциации с действием, направленным на устранение препятствия. Выражение *I dusted your ass* (Bad Boys for Life) употребляется в контексте дружеского соперничества между героями и означает: «Я тебя уделал». На уровне характеристики персонажа фраза подчеркивает уверенность в себе, конкурентность,

а также специфику речевого поведения героя, для которого важно подчеркнуть доминирование.

Синтаксический уровень

К данному уровню относятся устойчивые сленговые конструкции и междометия:

– *No way!* (Mean Girls) представляет собой короткое самостоятельное высказывание, передающее резкий протест или отказ. Его употребление демонстрирует импульсивность и эмоциональность героя: “Cady, you gotta steal that book. – No way!”

– *As if!* – выражение, которое функционирует как самостоятельный синтаксический элемент и представляет собой эллипсис — опущение частей полного высказывания. Полная форма конструкции могла бы звучать как “*As if I would ever do that!*”, однако в разговорной речи сохраняется только вводная часть, содержащая иронический или саркастический посыл. Подобный эллиптический оборот выступает в роли экспрессивного речения, выражающего насмешку, скепсис или недоверие. Использование данной фразы героиней подчеркивает её уверенность, и ироничную манеру общения.

– *What’s up, fellas?* (Bad Boys for Life) – фраза, которая представляет собой разговорное приветствие, характерное для неформального американского английского. Выражение *What’s up?* функционирует как синтаксическая формула, которая не требует буквального ответа о происходящем, а выступает в роли рутинного социального контакта, аналогичного «Привет» или «Как дела?». Обращение *fellas* — это сленговая форма слова *fellows*, традиционно используемая для обращения к группе друзей. Данное выражение демонстрирует привычную для героев фильма неформальность. Оно не несёт высокой смысловой нагрузки, но выполняет функцию установления контакта.

Подобные выражения придают ритмичность и эмоциональность высказыванию, что особенно важно для подростковых и молодежных фильмов.

Выводы

1. Молодёжный сленг в американских фильмах активно формируется за счёт различных словообразовательных моделей; фонетических искажений, редупликации, словосложения, усечения, аффиксации и метафоризации, — что делает речь персонажей более экспрессивной и приближённой к живому разговорному стилю.

2. Формирование сленговых выражений отражает не только языковые, но и социокультурные процессы, включая принадлежность к определенной социальной группе и желание выразить отношение к происходящему через эмоционально окрашенные и нестандартные формы.

3. Прагматическая функция сленгизмов в кинотексте заключается в усилении реалистичности диалогов, раскрытии характера героев и создании определённого коммуникативного эффекта.

Список литературы

1. Антрушина Г. Б., Афанасьева О. В. Лексикология английского языка. — М.: Дрофа, 2000. — 288 с.
2. Арсентьева Е. Ф. Редупликация в современном английском языке. — Казань: Казан. гос. ун-т, 2013. — 110 с.
3. Беляева Т. М. Нестандартная лексика английского языка. — Л.: Изд-во ЛГУ, 1985. — 120 с.
4. Берестовская Э. М. Молодежный сленг: структура и функции. — М., 1996. — 96 с.
5. Бирюков Н. Г. Молодежный сленг в призме фонетической мимикрии // Язык и культура. — 2021. — № 53 (1). — С. 50–63.
6. Вилюман В. Г. О способах образования слов сленга // Филологические науки. — 1960. — № 6. — С. 99–105.
7. Голуб И. Б. Современный русский язык. — М.: Логос, 2017. — 272 с.
8. Козина Т. А., Дидык Ю. А. Словообразовательные модели молодежного сленга // Молодежный сленг. — 2020. — № 2 (7). — С. 15–20.

9. Мелоян И. А. Основные способы образования сленга в современном английском языке // Филология и культура. — 2019. — № 3. — С. 45–50.
8. Резникова Л. В. Эволюция сленга в цифровую эпоху // Наука и образование. — 2022. — № 1. — С. 1–5.
10. Motion Picture Association. 2023 NTE Report MPA Comments [Электронный ресурс]. — 2023. — 39 с. URL: https://www.motionpictures.org/wp-content/uploads/2022/10/2023-NTE-Report-MPA-Comments.pdf (дата обращения: 05.05.2025).
11. Oxford English Dictionary. URL: https://www.oed.com/ (дата обращения: 05.05.2025).
12. Crystal, D. The Cambridge Encyclopedia of the English Language. 2nd ed. — Cambridge: Cambridge University Press, 2003. — 500 c.
13. Partridge, E. A Dictionary of Slang and Unconventional English. — London: Routledge, 1937. — 1360 c.
14. Partridge, E. Slang: Today and Yesterday. — London: Routledge, 1933. — 420 c.
15. Green, J. Green's Dictionary of Slang. — London: Chambers, 2010.
16. Thorne, T. Dictionary of Contemporary Slang. — London: Bloomsbury, 2006.

Сборник материалов

XIII Научно-практической конференции с международным участием

«НАУКА НАСТОЯЩЕГО И БУДУЩЕГО»

для студентов, аспирантов и молодых ученых

состоявшейся 15-17 мая 2025 г.

в г.Санкт-Петербурге

Том I