

На правах рукописи



ДЕМУШКИНА Ксения Михайловна

**МЕТОДИКИ И АЛГОРИТМЫ ОБРАБОТКИ ЦИФРОВЫХ ДАННЫХ
ЭЛЕКТРОКАРДИОГРАФИИ И БИОИМПЕДАНСОМЕТРИИ
ДЛЯ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ**

**2.3.1. Системный анализ, управление и обработка информации,
статистика (технические науки)**

А в т о р е ф е р а т
диссертации на соискание ученой степени
кандидата технических наук

ПЕНЗА – 2025

Работа выполнена на кафедре «Информационно-вычислительные системы» федерального государственного бюджетного образовательного учреждения высшего образования «Пензенский государственный университет».

Научный руководитель – доктор технических наук, доцент
Кузьмин Андрей Викторович

Официальные оппоненты: **Крошилин Александр Викторович**,
доктор технических наук, профессор,
федеральное государственное бюджетное
образовательное учреждение высшего
образования «Рязанский государственный
радиотехнический университет имени
В. Ф. Уткина» (г. Рязань), профессор
кафедры вычислительной и прикладной
математики;

Приоров Андрей Леонидович,
доктор технических наук, профессор,
федеральное государственное бюджетное
образовательное учреждение высшего
образования «Ярославский государственный
университет имени П. Г. Демидова»
(г. Ярославль), профессор кафедры
цифровых технологий и машинного
обучения

Ведущая организация – федеральное государственное бюджетное
образовательное учреждение высшего
образования «Тамбовский государственный
технический университет» (г. Тамбов)

Защита состоится 11 декабря 2025 г. в _____ часов на заседании диссертационного совета 24.2.357.03 на базе ФГБОУ ВО «Пензенский государственный университет» по адресу: 440026, г. Пенза, ул. Красная, 40.

С диссертацией можно ознакомиться в библиотеке и на сайте ФГБОУ ВО «Пензенский государственный университет»: https://dissov.pnzgu.ru/ecspertiza/Tehnicheskie_nauki/demushkina

Автореферат разослан «___» _____ 20__ г.

Ученый секретарь
диссертационного совета



Косников Юрий Николаевич

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Сердечно-сосудистые заболевания и онкологическая патология остаются наиболее распространенными в мире. Среди женщин наиболее часто встречается рак молочной железы. Это определяет высокую актуальность исследований, направленных на раннюю диагностику и прогнозирование сердечно-сосудистых и онкологических заболеваний.

Существует множество систем для диагностики заболеваний сердечно-сосудистой системы и рака молочной железы. В настоящее время подобные системы совершенствуются за счет внедрения методов искусственного интеллекта, анализа больших данных и мультимодальных подходов. Однако с увеличением сложности подобных систем растет и сложность интерпретации полученных результатов, особенно это актуально для систем, основанных на применении нейронных сетей, когда значимые диагностические параметры формируются и классифицируются самими сетями и часто не имеют медицинской интерпретации.

Результаты исследований в области обработки и анализа медицинских данных нашли отражение в работах таких отечественных ученых, как Б. А. Кобринский, В. А. Лишук, Т. В. Зарубина, А. В. Гусев, Н. Е. Шкловский-Корди, А. В. Крошилин, С. А. Филист, О. Н. Бодин, С. В. Фролов, В. Г. Полосин. Среди зарубежных ученых следует отметить L. T. Koczy, Edward H. Shortliffe, Kathryn A. Hannah, Wil van der Aalst, Boudewijn van Dongen, Massimiliano de Leoni, Marlon Dumas, Jan Mendling, которые внесли свой вклад в развитие медицинских экспертных систем, разработку методов анализа данных и др. Однако проблемы с интерпретацией данных по-прежнему остаются нерешенными, особенно в случае использования нейросетей. В связи с этим возникает задача разработки систем поддержки принятия решений, обладающих возможностью интерпретируемости. Под интерпретируемостью понимается представление получаемых решений в доступном для понимания виде с поэтапной визуализацией результатов. Это позволит повысить обоснованность решения и доверие к таким системам, что особенно важно в области медицины. Одним из путей решения данной актуальной задачи является применение технологии анализа процессов.

Объектом исследования является процесс принятия решений в области кардиографии и биоимпедансометрии.

Предметом исследования являются методики и алгоритмы обработки цифровых данных электрокардиографии и результатов биоимпедансометрии инструментами технологии анализа процессов и статистики.

Цель работы – совершенствование систем поддержки принятия врачебных решений в области анализа цифровых данных электрокардиографии и биоимпедансометрии.

Для достижения этой цели решаются следующие **задачи**:

1. Исследование методов и средств построения систем поддержки принятия решений для медицины с целью выявления эффективных методов обработки цифровых данных электрокардиографии и биоимпедансометрии, а также применения инструментов технологии анализа процессов.
2. Разработка методики анализа цифровых данных электрокардиографии с помощью инструментов технологии анализа процессов с целью улучшения интерпретируемости процесса анализа.
3. Разработка алгоритмов для формирования цифровых данных электрокардиографии в виде журнала событий с целью дальнейшей обработки данных инструментами технологии анализа процессов.
4. Разработка методики анализа цифровых данных биоимпедансометрии для комплексной оценки различий параметров объектов и представления данных с помощью инструментов технологии анализа процессов.
5. Экспериментальное подтверждение работоспособности предложенных методик и алгоритмов.

Методы исследования. Основными методами исследования являются компьютерная обработка информации, анализ процессов, теория принятия решений, цифровая обработка сигналов, биоимпедансная спектроскопия, статистический регрессионный и контурный анализ.

Научная новизна:

1. Разработана методика анализа цифровых данных электрокардиографии, отличающаяся применением технологии анализа процессов, что позволяет улучшить интерпретируемость процесса анализа путем контроля промежуточных результатов за счет представления процесса сердечной деятельности в виде журнала событий и последующей компьютерной обработки и визуализации этапов поиска решения (п. 2, 12 специальности 2.3.1).
2. Предложен модифицированный алгоритм определения R-пика на цифровых данных электрокардиографии, основанный на алгоритме Пана – Томпкинса, отличающийся адаптивным определением границ QRS-комплекса, который позволяет повысить точность определения пиков в среднем на 2 % во всех отведениях электрокардиографии по сравнению с аналогами (п. 4 специальности 2.3.1).
3. Предложен алгоритм формирования цифровых данных электрокардиографии для визуализации процессов сердечной деятельности, отличающийся тем, что данные представляются в виде журнала событий, что позволяет обеспечить возможность анализа цифровых данных электрокардиографии средствами технологии анализа процессов и повысить интерпретируемость процесса принятия решения (п. 4, 12 специальности 2.3.1).
4. Предложена методика анализа данных биоимпедансометрии, отличающаяся применением комплексной оценки статистических параметров и метрики Хаусдорфа, что позволяет ранжировать объект по возрастанию неоднородных включений и расширить возможности практического применения метода биоимпедансометрии (п. 4 специальности 2.3.1).

Теоретическая и практическая значимость. Автором впервые предложена методика анализа цифровых данных электрокардиографии (ЭКГ) средствами технологии анализа процессов, которая позволяет осуществлять анализ биомедицинских данных (ЭКГ) за счет визуализации процесса, описанного ими. Такая методика позволит использовать технологию анализа процессов в системах поддержки принятия решений с использованием описания сердечной деятельности в виде журнала событий.

Работа обобщает проведенные автором теоретические и экспериментальные исследования, связанные с алгоритмами обработки цифровых данных электрокардиографии и биоимпедансометрии. Разработанный модифицированный алгоритм Пана – Томпкинса обеспечивает высокую точность (74 %) в сравнении с аналогами (классический алгоритм Пана –Томпкинса, алгоритм Neurokit), что позволяет улучшить качество идентификации R-пику в зашумленных сигналах ЭКГ, повышая эффективность автоматизированного анализа цифровых данных электрокардиографии.

Разработанный алгоритм сегментации и формирования цифровых данных электрокардиографии позволяет использовать преимущества технологии анализа процессов, таких как визуализация потоков событий, выявление отклонений и аномалий, что в итоге обеспечивает специалистам возможность детально исследовать электрокардиографические данные.

Работа выполнялась в рамках следующих финансируемых проектов:

1. Государственное задание Министерства науки и высшего образования Российской Федерации «Разработка технологии раннего обнаружения новообразований молочной железы на основе методов микроволновой томографии и биоимпедансной спектроскопии» (рег. № 124020200015-7), Пензенский государственный университет.

2. Грант «Ректорские гранты для молодых ученых и аспирантов – 2023» Пензенского государственного университета (приказ № 78/О от 25.01.2023). «Разработка и исследование способа и программного обеспечения для анализа ЭКГ на основе подхода Process Mining с возможностью построения деревьев диагностических решений».

Положения, выносимые на защиту:

1. Методика анализа цифровых данных электрокардиографии с помощью технологии анализа процессов, позволяющая улучшить интерпретируемость процесса анализа.

2. Модифицированный алгоритм определения R-пики, основанный на алгоритме Пана – Томпкинса, обеспечивающий повышение точности определения пиков в среднем на 2 % во всех отведениях электрокардиографии по сравнению с аналогами.

3. Алгоритм формирования цифровых данных электрокардиографии для анализа и визуализации процессов сердечной деятельности, позволяющий использовать инструменты технологии анализа процессов для оценки состояния сердца.

4. Методика анализа данных биоимпедансометрии на основе комплексной оценки статистических параметров и метрики Хаусдорфа, обеспечивающая возможность практического применения метода биоимпедансометрии.

Степень достоверности и апробация работы. Основные результаты диссертационной работы были представлены и обсуждены на научных конференциях международного и всероссийского уровня: XXVII Международная конференция «Цифровая обработка сигналов и ее применение – DSPA–2025» (Институт проблем управления РАН, г. Москва, 2025 г.); научно-техническая конференция с международным участием «Информационные технологии и системы (ИСТ–2023)» (г. Самара, 2023 г.); XXIII Международная научно-техническая конференция «Проблемы информатики в образовании, управлении, экономике и технике» (Пензенский государственный университет, г. Пенза, 2023 г.); Международная конференция «Мерно-информационные технологии» MIT–2023 (г. Нови-Сад, Сербия); X Всероссийская научно-практическая конференция «Информационные технологии в науке и образовании. Проблемы и перспективы» (Пензенский государственный университет, г. Пенза, 2023 г.); Всероссийская научно-практическая конференция «Студент-Наука» (г. Воронеж, 2022 г.); университетская научно-практическая конференция «Актуальные проблемы науки и образования» (Пензенский государственный университет, г. Пенза, 2022–2024 гг.). Достоверность подтверждается внедрением результатов в научную и практическую деятельность ряда организаций, а также апробацией на международных и всероссийских конференциях.

Соответствие паспорту специальности. Диссертационная работа соответствует паспорту специальности 2.3.1. Системный анализ, управление и обработка информации, статистика по следующим пунктам: п. 2 – «Формализация и постановка задачи системного анализа, оптимизации, управления, принятия решения, обработки информации и искусственного интеллекта»; п. 4 – «Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта»; п. 12 – «Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации».

Публикации. Основные положения и результаты выполненных исследований отражены в 15 публикациях, из них 5 статей в изданиях, рекомендуемых ВАК при Минобрнауки России; 2 публикации в изданиях, индексируемых в международной базе данных SCOPUS; 5 публикаций в российских журналах, материалах международных, всероссийских, региональных научно-практических конференций; 3 свидетельства о государственной регистрации программ для ЭВМ.

Внедрение результатов работы. Предложенные методы и алгоритмы обработки цифровых данных электрокардиографии и биоимпедансометрии используются в разработках и исследованиях:

1) Пензенского государственного университета в рамках государственного задания Министерства науки и высшего образования Российской Феде-

рации «Разработка технологии раннего обнаружения новообразований молочной железы на основе методов микроволновой томографии и биоимпедансной спектроскопии» (рег. № 124020200015-7);

2) Пензенского государственного университета в рамках учебного процесса по направлениям подготовки 09.03.03 «Прикладная информатика» и 09.05.01 «Применение и эксплуатация автоматизированных систем специального назначения»;

3) АО «Научно-производственное предприятие "Рубин"» (г. Пенза) при разработке проекта использования технологии анализа процессов на мобильных станциях забора крови;

4) ООО «Максофт» (г. Москва) при разработке проектов по выявлению уязвимостей в библиотеках исходного кода (MaSCA) с помощью методики применения плагина Multi-perspective Process Explorer фреймворка ProM, а также при разработке проекта «Школа Максимум» использовалась методика применения плагина Dotted Chart и Filter Log by attribute фреймворка ProM;

5) ФГБОУ ВО «Самарский государственный медицинский университет Минздрава России» (г. Самара) при подготовке специалистов Передовой медицинской инженерной школы по программам «Инженерия искусственного интеллекта» и «Телемедицина и поддержка принятия врачебных решений» направления подготовки 09.04.01 «Информатика и вычислительная техника».

Личный вклад автора. Основные результаты, выносимые на защиту, получены автором лично. Результаты, опубликованные совместно с другими авторами, принадлежат авторам в равных долях. Результаты других авторов, которые использованы при изложении результатов работы, снабжены ссылками на соответствующие источники. За высокий уровень научно-технических и инженерных решений, применяемых в программных разработках в рамках диссертации, соискатель удостоен звания лауреата регионального конкурса «Инженер года» в номинации «Инженер-программист» (г. Пенза, 2022 г.).

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка литературы, включающего 142 источника, и 2 приложений. Основная часть работы изложена на 152 страницах, содержит 87 рисунков и 17 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении содержатся обоснование актуальности и научной новизны диссертации, цель и задачи исследования, результаты, выносимые на защиту, научно-практическая значимость, сведения об апробации работы и личном вкладе автора.

Первая глава посвящена обзору и анализу методов и алгоритмов обработки цифровых данных электрокардиографии и биоимпедансометрии для поддержки принятия решений. Рассмотрены существующие методы проекти-

рования систем поддержки принятия решений, выделены недостатки и достоинства. Наиболее популярными на сегодня являются мультимодальные системы поддержки принятия решений за счет комплексного принятия решения на основе разнородных данных. Проведен анализ существующих методик и алгоритмов анализа цифровых данных электрокардиографии и биоимпедансометрии. Определены недостатки существующих средств анализа цифровых данных электрокардиографии и биоимпедансометрии. На основе проведенного анализа обоснована актуальность работы, сформулированы цель и задачи исследования.

Вторая глава посвящена разработке методики анализа цифровых данных средствами технологии анализа процессов, а также разработке необходимых алгоритмов для обработки информации. Технология анализа процессов – это технология, обеспечивающая извлечение информации о ходе процесса из журнала событий с целью дальнейшего его прогнозирования и усовершенствования. Под процессом в данном случае понимается последовательность взаимосвязанных действий (событий), выполняемых для достижения определенной цели. Журнал событий – это запись всех событий, которые происходят в процессе, с указанием времени появления данного события в формате XES. XES – это стандарт для журналов событий, основанный на XML, необходимый для работы с инструментами технологии анализа процессов. Используя данные журнала, можно построить модель процесса, провести сравнение с другой моделью процесса, выявив уязвимости в действующей модели, осуществить прогнозирование поведения модели при различных входных параметрах. Достоинством технологии анализа процессов является возможность объективного выявления реального хода процессов, включая скрытые зависимости и отклонения от нормативных моделей, что позволяет повысить прозрачность деятельности и обоснованность принимаемых решений. Основная идея данной методики заключается в формировании цифровых данных электрокардиографии в виде журнала событий с целью последующего анализа средствами технологии анализа процессов (рисунок 1).

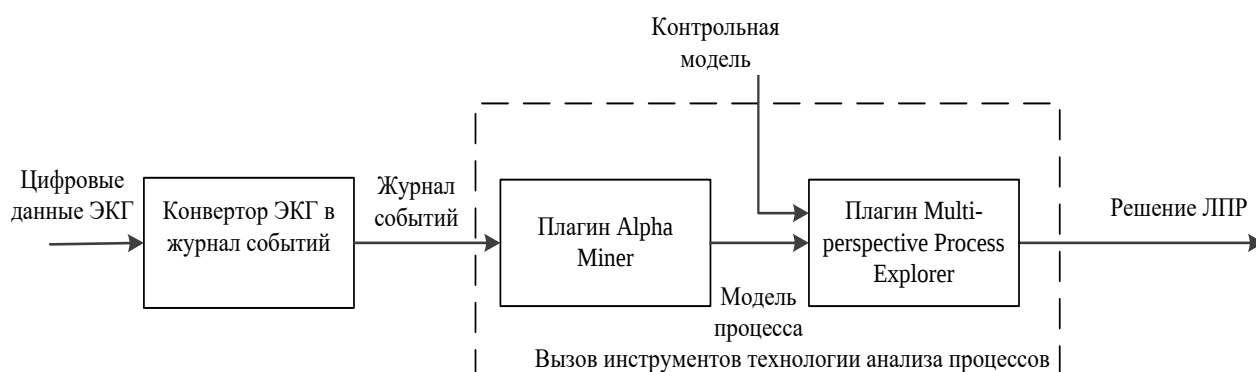


Рисунок 1 – Схема анализа цифровых данных ЭКГ средствами технологии анализа процессов

Разработанная методика состоит из следующих этапов:

- 1) трансформация цифровых данных электрокардиографии в журнал событий;
- 2) построение модели процесса в виде сети Петри с помощью плагина Alpha Miner фреймворка ProM;
- 3) сравнение контрольной модели и модели исследуемой электрокардиографии с помощью плагина Multi-perspective Process Explorer фреймворка ProM;
- 4) принятие решения на основе анализа отклонений исследуемой модели электрокардиографии от контрольной, определения продолжительности пиков с помощью графического интерфейса плагина Multi-perspective Process Explorer фреймворка ProM.

Основная задача Multi-perspective Process Explorer заключается в сопоставлении контрольной модели процесса с вводимым пользователем журналом событий (рисунок 2). Справа расположена информационная панель (information panel), которая представляет собой статистику по имеющимся и пропущенным событиям, а также расчетное значение соответствия исследуемой модели контрольной (Avg fitness). В верхней части интерфейса расположена сеть Петри с цветовым выделением пропущенных событий (P+start, P+complete) и средним временем выполнения каждого события.

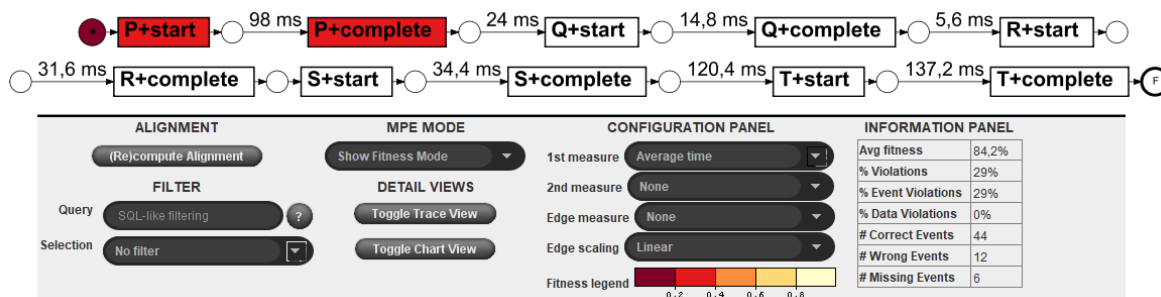


Рисунок 2 – Интерфейс плагина MPPE

Плагин MPPE предоставляет возможность исследования каждого кардиоцикла в отдельности (рисунок 3), подсвечивая пропущенные события (фиолетовый блок) и неизвестные события (желтый блок).

Таким образом, разработанная методика позволяет повысить наглядность процесса сердечной деятельности. Плагин Multi-perspective Process Explorer дает возможность в процессе сопоставления эталонной модели процесса с вводимым пользователем журналом событий:

- 1) определить среднее/минимальное/максимальное время длительности;
- 2) оценить общую модель исследуемого кардиосигнала;
- 3) рассмотреть модель каждого кардиоцикла в отдельности;
- 4) выделить не соответствующие нормам пики кардиоцикла;
- 5) предоставить доступ пользователю к амплитудному значению каждого кардиопика.

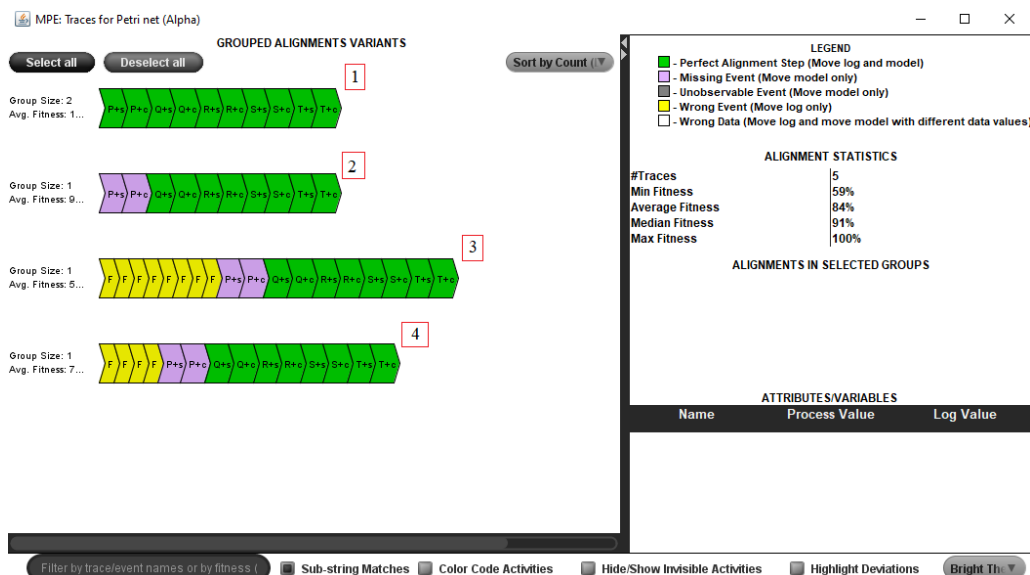


Рисунок 3 – Интерфейс плагина МРРЕ (по группам)

Для реализации рассмотренной методики необходима трансформация цифровых данных электрокардиографии в журнал событий. На рисунке 4 представлена общая схема алгоритма формирования цифровых данных кардиограммы в виде журнала событий формата хес.



Рисунок 4 – Схема алгоритма трансформации сегментированных цифровых данных электрокардиографии в журнал событий

Алгоритм включает предложенную автором модификацию алгоритма Пана – Томпкинса (символ 2), в основу которой положено динамическое определение расстояния между пиками после процедуры фильтрации:

$$\text{distance} = 0.2 f + 10, \quad (1)$$

где f – частота дискретизации сигнала (Гц); 0.2 – длительность QRS-комплекса (с); 10 – константа, учитывающая сдвиг сигнала при фильтрации.

Для уточнения местоположения пика на исходном сигнале автором предлагается рассчитать интервал предполагаемого R-пика по формуле

$$\text{width_R} = 0.5tf, \quad (2)$$

где f – частота дискретизации сигнала (Гц); t – длительность R-пика, равная 0.1 с; 0.5 – экспериментально полученный коэффициент для уменьшения погрешности определения пика.

Результатом работы алгоритма является массив значений отсчетов сигнала электрокардиографии, где найден R-пик.

Алгоритм сегментации позволяет разбить сигнал на характерные пики ЭКГ – P, Q, R, S, T, определяя положение этих пиков, а также время их начала и окончания. Для задачи определения начала и окончания пика берется точка пересечения сигнала с изолинией.

Разработанный алгоритм сегментации позволяет представить цифровые данные электрокардиографии в виде журнала событий формата XES, поскольку временные метки начала и окончания пиков известны, а также определены их амплитудные значения.

Таким образом, разработанная методика обеспечивает визуализацию процесса анализа цифровых данных электрокардиографии средствами технологии анализа процессов. Представленная методика обеспечивает принцип «белого ящика», что позволяет сделать процесс анализа цифровых данных ЭКГ интерпретируемым для лица, принимающего решение. Для реализации методики с целью повышения ее точности модифицирован алгоритм Пана – Томпкинса. Разработанный алгоритм сегментации позволил выделить из цифровых данных электрокардиографии начало и окончание всех пиков, что дало возможность формирования информации в виде журнала событий формата XES.

В третьей главе предложена методика анализа цифровых данных биоимпедансометрии для систем поддержки принятия решений с помощью технологии анализа процессов. Автором предлагается использовать биоимпедансометрию для диагностики рака молочной железы за счет аппаратно-программного комплекса, а также однородных и неоднородных фантомов в качестве объекта исследования.

Результатом работы аппаратно-программного комплекса является спектр диэлектрической проницаемости, полученный методом биоимпедансной спектроскопии, который оценивается с помощью следующих статистиче-

ских параметров: среднеквадратичной ошибки (MSE), среднеквадратичного отклонения (RMSE), средней абсолютной ошибки (MAE), средней абсолютной ошибки, выраженной в процентах (MAEP), и метрики Хаусдорфа. Чем больше отклонение исследуемого объекта от контрольного, тем больше объем неоднородного включения. В качестве метода нормирования был выбран метод MinMax, поскольку он обеспечивает приведение данных к единому масштабу [0; 1], где минимальные значения отображаются вблизи 0, а максимальные – вблизи 1.

Для агрегации предлагается использовать метод главных компонент, который позволяет уменьшить размерность выборки без потери информативности. Предварительно с помощью программных средств рассчитываются матрица ковариации, собственные значения и собственный вектор, затем выделяется главная компонента и строится вектор признаков. Выражение (3) представляет собой матрично-векторную запись процедуры: нормализация признаков в пределах [0; 1], центрирование относительно среднего вектора μ и проекция на вектор (или матрицу) весов W , определяющих направление главных компонент. В результате формируется обобщенная числовая оценка E отличия спектра от контрольного объекта:

$$E = W \left(\frac{x - \min(x)}{\max(x) - \min(x)} - \mu \right), \quad (3)$$

где x – вектор признаков; $\min(x)$, $\max(x)$ – векторы минимумов и максимумов; μ – вектор средних значений нормализованных признаков; W – вектор весов признаков.

Для визуализации полученных результатов используются инструменты технологии анализа процессов, в частности плагины Dotted Chart и Filter Log by attribute фреймворка ProM (рисунок 5).

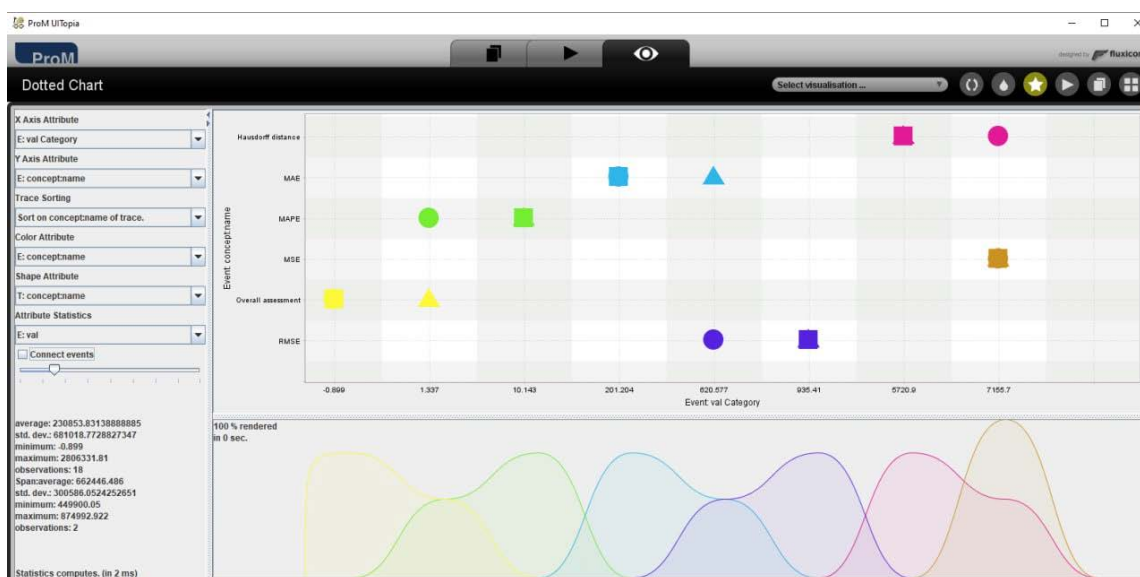


Рисунок 5 – Результат работы плагина Dotted Chart

По оси абсцисс указывается значение рассматриваемых метрик, по оси ординат – расчетные метрики и их агрегированное значение (Overall assessment). Значение одной метрики обозначается цветом, а принадлежность метрик к конкретной дате исследования – различными геометрическими фигурами. Такая графическая визуализация позволяет рассмотреть градацию метрик по их значениям от исследования к исследованию, а инструменты плагина Dotted Chart дадут возможность при необходимости изменить значение оси для анализа данных по времени или в пределах исследования.

Разработана следующая методика для анализа цифровых данных биоимпедансометрии для систем поддержки принятия решений с помощью технологии анализа процессов:

Этап 1. Получение данных активного и реактивного сопротивления исследуемого объекта с помощью RLC-метра в режиме List Sweep.

Этап 2. Расчет статистических параметров и метрики Хаусдорфа для всех исследуемых объектов относительно контрольного объекта.

Этап 3. Приведение рассматриваемых параметров к одной единице измерения с помощью метода нормирования MinMax.

Этап 4. Приведение рассматриваемых параметров к одной оценке по каждому исследуемому объекту с помощью метода агрегации PCA.

Этап 5. Визуализация статистических параметров, метрики Хаусдорфа и агрегированной оценки с помощью плагинов Dotted Chart и Filter Log by attribute фреймворка ProM.

Предложенная методика анализа данных биоимпедансной спектроскопии на основе комплексной оценки с использованием статистических параметров и метрики Хаусдорфа с помощью инструментов технологии анализа процессов позволяет количественно оценить различия спектра диэлектрической проницаемости, полученного методом биоимпедансной спектроскопии.

В четвертой главе представлена схема взаимодействия разработанных методик, использующих технологии анализа процессов, с системами поддержки принятия решений, а также описаны программные реализации разработанных во второй и третьей главах алгоритмов.

Взаимодействие инструментов технологии анализа процессов осуществляется за счет headless-запуска. Headless-Launcher запускает фреймворк ProM без использования GUI, позволяя запустить фреймворк на сервере. Headless-launcher создается за счет сборки из исходных файлов или за счет написания headless-launcher через JavaAPI. Использование headless-launcher позволит обеспечить взаимодействие инструментов технологии анализа процессов с любыми системами поддержки принятия решений, тем самым обеспечив мультимодальность за счет возможности обработки разнородных данных (журнал событий) (рисунок 6).

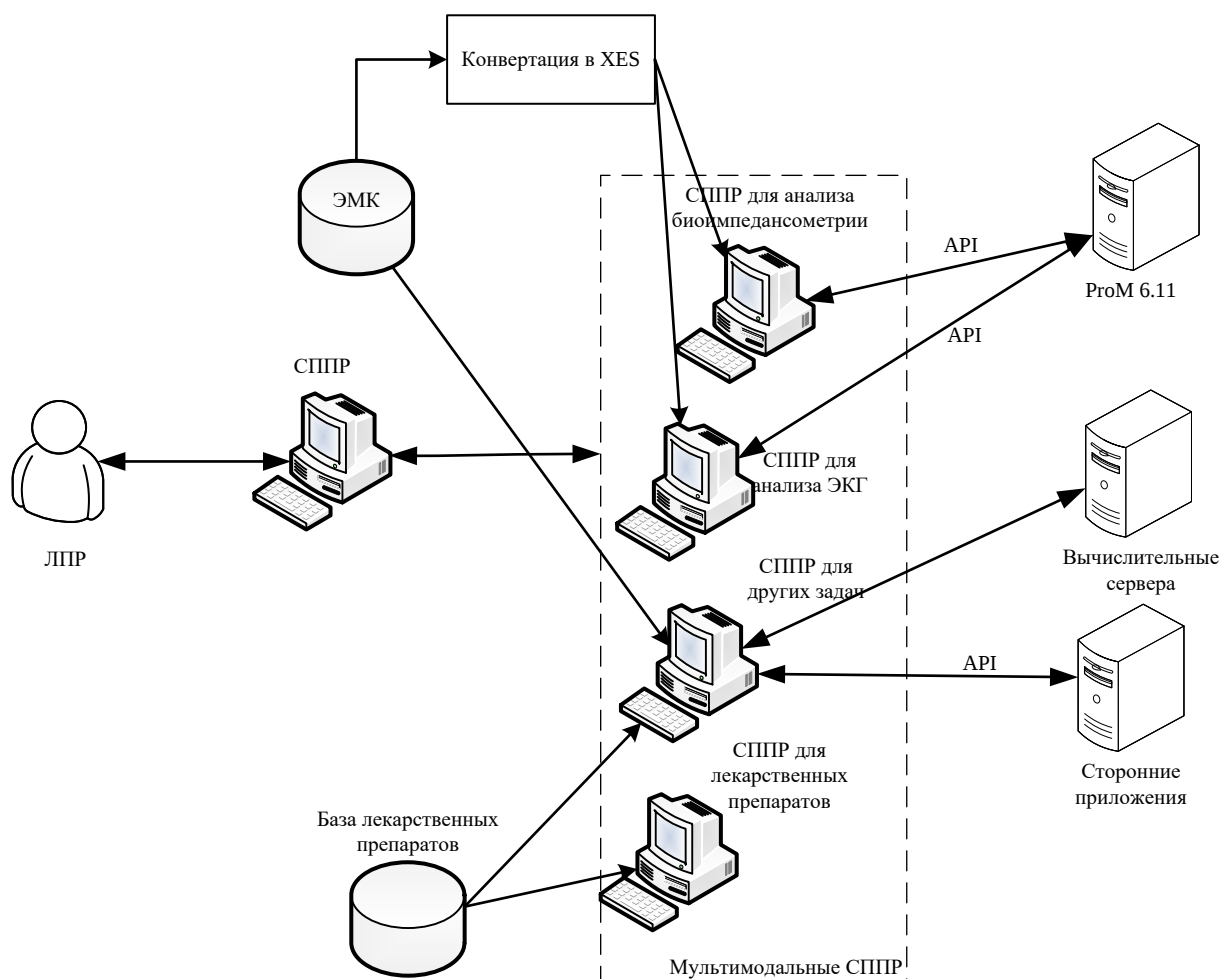


Рисунок 6 – Схема взаимодействия разработанных методик с системами поддержки принятия решений

Программы разработаны на языке программирования Python с использованием следующих библиотек: numpy, SciPy, PM4Py, hausdorff, sklearn.metrics, klearn.decomposition и sklearn.preprocessing.

В главе приведены экспериментальные подтверждения эффективности разработанных алгоритмов на открытой библиотеке кардиосигналов PhysioNet. Проведен сравнительный анализ модифицированного алгоритма Пана – Томпкинса и алгоритма сегментации, который является частью алгоритма трансформации цифровых данных кардиографии в журнал событий. Модифицированный алгоритм показывает высокие значения чувствительности (99 %), специфичности (90 %), точности (91 %) по сравнению с аналогами. Алгоритм сегментации сравнивался с алгоритмом определения пиков в библиотеке Neurokit языка программирования Python. Результаты сравнительного анализа показали высокую точность определения каждого пика с помощью авторского алгоритма сегментации кардиосигнала, несмотря на шумы и артефакты. Также важно отметить, что, в отличие от аналога, авторский алгоритм сегментации способен определять начало и окончание каждого из пиков, что

является важной информацией для представления цифровых данных ЭКГ в виде журнала событий.

Алгоритм анализа цифровых данных биоимпедансометрии представляет собой программу, которая обеспечивает расчет статистических параметров и метрики Хаусдорфа. Используемые библиотеки дают возможность нормировать значения в пределах $[0; 1]$, а также агрегировать с помощью метода главных компонент. Полученные оценки позволяют ранжировать объекты по возрастанию объема неоднородности.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В диссертационной работе на основании выполненных автором исследований решена научная задача обоснования, разработки и исследования комплекса решений, обеспечивающих создание методик и алгоритмов анализа цифровых данных электрокардиографии и биоимпедансометрии для систем поддержки принятия решений:

1. Предложена методика анализа цифровых данных электрокардиографии с помощью технологии анализа процессов, что дало возможность улучшить интерпретируемость поиска решения для специалиста за счет визуализации процессов, зарегистрированных в журнале событий, за счет разработанных плагинов, предоставляя пользователям интерфейс для взаимодействия. Технология анализа процессов обеспечивает анализ данных методом «белого ящика», что обеспечивает высокую интерпретируемость процесса анализа сердечной деятельности, в отличие от метода «черного ящика».

2. На основе алгоритма Пана – Томпкинса разработан модифицированный алгоритм для определения *R*-пика с адаптивным определением границ *QRS*-комплекса, что обеспечивает более точную идентификацию пика. Это позволяет повысить точность определения пиков в среднем на 2 % во всех отведениях ЭКГ по сравнению с аналогами. Наибольшую точность разработанный алгоритм продемонстрировал на отведении *V6*, она составила 91 %, чувствительность составила 99 %. Это говорит об уверенном обнаружении пика, что влияет на качество автоматизированного анализа цифровых данных ЭКГ.

3. Разработан алгоритм формирования цифровых данных электрокардиографии в виде журнала событий стандарта XES, что дает возможность анализировать данные средствами технологии анализа процессов, обеспечить визуализацию процесса сердечной деятельности, интерактивно взаимодействовать с инструментами анализа. В частности, использование плагина Multi-perspective Process Explorer фреймворка ProM 6.11 обеспечивает визуальное выделение пропущенных или не соответствующих эталонной модели событий, что наглядно отражает аномалии в данных, описывающих сердечную деятельность.

4. Предложена методика анализа данных биоимпедансной спектроскопии на основе комплексной оценки статистических параметров и метрики

Хаусдорфа, позволяющая оценить различия кривых амплитудно-частотных характеристик объекта, что обеспечивает возможность ранжирования исследуемых объектов по степени изменения биомедицинских характеристик. Применение технологии анализа процессов позволяет представить данные биоимпедансной спектроскопии в визуальном интерпретируемом виде.

5. Программная реализация результатов на языке Python с использованием библиотек, фреймворка, плагинов и т.п. подтверждает работоспособность предложенных решений.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

*Статьи в изданиях, рекомендованных ВАК при Минобрнауки России
по специальности 2.3.1*

1. Демушкина, К. М. Обзор методов проектирования систем поддержки принятия врачебных решений / К. М. Демушкина, М. О. Демушкин, А. В. Кузьмин // Модели, системы, сети в экономике, технике, природе и обществе. – 2022. – № 2 (42). – С. 75–89.

2. Демушкина, К. М. Анализ возможностей инструментов реализации технологии process mining / К. М. Демушкина, А. В. Кузьмин // Известия Самарского научного центра Российской академии наук. – 2023. – Т. 25, № 4 (114). – С. 114–120.

3. Демушкина, К. М. Применение плагина Multi-perspective Process Explorer программы Prom 6.11 для анализа кардиограммы / К. М. Демушкина, А. В. Кузьмин // Известия высших учебных заведений. Приборостроение. – 2025. – Т. 68, № 5. – С. 427–437. – doi: 10.17586/0021-3454-2025-68-5-427-437

4. Демушкина, К. М. Модифицированный алгоритм Пана-Томпкинса с адаптивным определением границ QRS-комплекса / К. М. Демушкина, А. В. Кузьмин // Цифровая обработка сигналов. – 2025. – № 2. – С. 75–80.

5. Демушкина, К. М. Агрегированная оценка результатов диэлектрической импедансной спектроскопии на основе статистических параметров и метрики Хаусдорфа / К. М. Демушкина // Модели, системы, сети в экономике, технике, природе и обществе. – 2025. – № 2. – С. 147–155.

Публикации в изданиях, индексируемых в наукометрической базе Scopus

6. Demushkina, K. Method and Process Mining Based Software for ECG Analysis and Arrhythmia Detection / K. Demushkina, A. Kuzmin. – М. : IEEE. – 2025. – P. 1–5.

7. Baranov, V. Estimation of the Parameters of the Havriliak-Negami Model for Breast Phantoms / V. Baranov, A. Kuzmin, Yu. Smirnov, V. Martynova, K. Demushkina – М. : IEEE, 2025. – P. 1–5.

8. Демушкина, К. М. Исследование возможности построения дерева принятия решения при постановке диагноза с использованием ProM / К. М. Демушкина // Студент-Наука : сборник трудов Всероссийской научно-практической конференции (г. Воронеж, 5–6 декабря 2022 г.). – Воронеж : Воронежский государственный технический университет, 2022. – С. 106–108.

9. Демушкина, К. М. Разработка плагина для работы с цифровыми данными ЭКГ в системе ProM 6.11 / К. М. Демушкина, А. В. Кузьмин // Информационные технологии в науке и образовании. Проблемы и перспективы : сборник статей по материалам X Всероссийской научно-практической конференции (г. Пенза, 15 марта 2023 г.). – Пенза : Изд-во ПГУ, 2023. – С. 167–170.

10. Демушкина, К. М. Особенности разработки плагина для анализа биомедицинских данных в среде ProM 6.11 / К. М. Демушкина, А. В. Кузьмин // Информационные системы и технологии (ИСТ–2023) : труды научно-технической конференции с международным участием (г. Самара, 19–21 июня 2023 г.). – Самара : Самарский научный центр РАН, 2023. – С. 199–202.

11. Демушкина, К. М. Особенности реализации плагина для построения дерева решений в среде ProM / К. М. Демушкина // Актуальные проблемы науки и образования : сборник научных статей по материалам Всероссийской научно-практической конференции профессорско-преподавательского состава, аспирантов и студентов, посвященной 80-летию Пензенского государственного университета (г. Пенза, 19–25 октября 2023 г.). – Пенза : Изд-во ПГУ, 2023. – С. 11–14.

12. Демушкина, К. М. Разработка пользовательского конфигурационного окна для анализа ЭКГ в среде ProM / К. М. Демушкина, А. А. Масленников // Проблемы информатики в образовании, управлении, экономике и технике : сборник статей XXIII Международной научно-технической конференции, посвященной 80-летию Пензенского государственного университета (Пенза, 24–25 ноября 2023 г.). – Пенза : Изд-во ПГУ, 2023. – С. 158–161.

Свидетельства о регистрации программ для ЭВМ

13. Свидетельство о государственной регистрации программы для ЭВМ № 2023680277 Российская Федерация. Плагин для определения сердечного ритма на основе данных ЭКГ для системы ProM 6.11 : № 2023668910 : заявл. 14.09.2023 : опубл. 28.09.2023 / А. В. Кузьмин, К. М. Демушкина ; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Пензенский государственный университет».

14. Свидетельство о государственной регистрации программы для ЭВМ № 2025616658 Российская Федерация. Программа обнаружения R-пика на сигнале ЭКГ с адаптивным определением границ диапазона его расположения : № 2025615149 : заявл. 14.03.2025 : опубл. 19.03.2025 / А. В. Кузьмин,

К. М. Демушкина ; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Пензенский государственный университет».

15. Свидетельство о государственной регистрации программы для ЭВМ № 2025665861 Российская Федерация. Программа для сравнительного анализа результатов биоимпедансной спектроскопии с использованием статистических показателей и метрики Хаусдорфа : № 2025663669 : заявл. 30.05.2025 : опубл. 19.06.2025 / А. В. Кузьмин, К. М. Демушкина ; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Пензенский государственный университет».

Научное издание

ДЕМУШКИНА Ксения Михайловна

**МЕТОДИКИ И АЛГОРИТМЫ ОБРАБОТКИ ЦИФРОВЫХ ДАННЫХ
ЭЛЕКТРОКАРДИОГРАФИИ И БИОИМПЕДАНСОМЕТРИИ
ДЛЯ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ**

2.3.1. Системный анализ, управление и обработка информации,
статистика (технические науки)

Редактор *А. А. Есавкина*

Технический редактор *Ю. В. Анурова*

Компьютерная верстка *Ю. В. Ануровой*

Подписано в печать 10.10.2025. Формат 60×84¹/₁₆.

Усл. печ. л. 1,16. Заказ № 498. Тираж 100.

Издательство ПГУ.

440026, г. Пенза, ул. Красная, 40.

Тел./факс: (8412) 66-60-49, 66-67-77; e-mail: iic@pnzgu.ru