

ПЕРСОНАЛЬНЫЕ ПРОФИЛИ ПАЦИЕНТОВ С БРОНХОЛЕГОЧНЫМИ ЗАБОЛЕВАНИЯМИ

М. Р. Богданов • Р. В. Насыров • Г. Р. Шахмамметова

Аннотация. Проведены исследования в области разработки персональных профилей пациентов с бронхолегочными заболеваниями. Для определения компонентов профиля было использовано три подхода анализа дыхательных шумов, два из которых были основаны на ортогональных вейвлетах, а один — на спектральных характеристиках дыхательных шумов. Было установлено, что ортогональные вейвлеты обладают высокой селективностью к различным физиологическим состояниям органов дыхания. Оценка эффективности классификации бронхолегочных заболеваний, когда в качестве признаков использовались вейвлеты показала, что высокой обобщающей способностью обладают такие методы, как случайный лес и экстремальный градиентный бустинг. Аналогичный результат был получен при использовании в качестве признаков таких спектральных характеристик дыхательных шумов как среднеквадратичная энергия (rmse), спектрограмма мощности, спектральный центроид, спектральная полоса пропускания p -го порядка, частота спада, скорость пересечения нуля временного аудио ряда, мелкочастотные кепстральные коэффициенты (MFCC). Предложенные подходы делают возможным формирование компонентов персонального профиля пациента на основе скрининга пациентов с бронхолегочными заболеваниями.

Ключевые слова: персональный профиль пациента; бронхолегочные заболевания; машинное обучение; вейвлет-анализ.

ВВЕДЕНИЕ

В настоящее время большую актуальность приобретает внедрение в медицинскую практику «четырёх Р»:

- Prediction — предупреждение заболеваний;
- Prevention — предотвращение заболеваний;
- Personalization — персонализированный подход к пациенту;
- Participation — участие и полное понимание процессов пациентом.

Персонализация в различных работах трактуется достаточно широко от создания профиля пациента, содержащего информацию о поле, возрасте, расе, финансовом положении, данных о перенесенных ранее заболеваниях [Cre25] до генетического профайлинга [Sar19]. Учет индивидуальных особенностей пациентов повышает точность прогноза течения болезни, позволяет повысить эффективность лечения. При составлении персональных профилей пациентов важны как социальные, так и биологические данные. Так, например, выяснилось, что в соседних округах одного и того же штата смертность от COVID-19 может различаться в разы. Дело в том, что в одном округе преимущественно живут белые американцы с высоким уровнем дохода, в то время как в соседнем округе проживают преимущественно черные американцы, не имеющие медицинской страховки. Широкую известность получило американское племя пима [Smi88, Sch06], в котором в следствие инбридинга велик уровень наследственных заболеваний, в частности диабета.

Цифровой медицинский профиль (цифровой двойник) – это совокупность базовых, демографических, клинических, генетических и медико-финансовых (ОМС/ДМС) данных о пациенте, позволяющих сформировать единую картину о состоянии здоровья пациента и оказанных ему медицинских услугах с базовым набором пациентских сервисов и сервисов доступа (<https://minzdrav.gov.ru>).

Данное исследование посвящено разработке персонализированных подходов оценки состояния пациентов с бронхолегочными заболеваниями (БЛЗ) и формирования предлагаемых

профилей пациентов с БЛЗ. Эти заболевания являются третьей по частоте причиной смертности во всем мире [Roc], они ложатся тяжким бременем на систему общественного здравоохранения [WHO17]. Для успешного лечения респираторных заболеваний очень важна ранняя диагностика [Mar14]. Желательно, чтобы диагностический метод позволял проводить массовый скрининг подобно флюорографии, позволяя быстро отделять здоровых испытуемых от лиц с различными бронхолегочными патологиями.

При диагностике респираторных заболеваний широко применяется аускультация, основанная на анализе дыхательных шумов [Ear00]. Дыхательные шумы являются важными индикаторами состояния органов дыхания. Звук, издаваемый при дыхании человека, напрямую связан с движением воздуха, изменениями в легочной ткани и положением выделений в легких. Свистящий звук, например, является признаком того, что у пациента есть обструктивное заболевание дыхательных путей, такое как астма или хроническая обструктивная болезнь легких (ХОБЛ) [Pii95, Sar15, Sov00, Pra17, Cha16]. Автоматизированная классификация дыхательных шумов может выявлять аномалии на ранних стадиях дыхательной дисфункции и, таким образом, повышать эффективность принятия решений. Появление цифровых стетоскопов облегчает использование телемедицинских технологий для такой автоматизированной диагностики.

ПОСТАНОВКА ЗАДАЧИ

Целью настоящего исследования является разработка методов создания персональных профилей пациентов с бронхолегочными заболеваниями на основе анализа дыхательных шумов.

Для достижения цели исследования необходимо решить следующие задачи. Во-первых, определить характерные частоты дыхательных шумов. Во-вторых, построить компоненты профиля пациента в форме эмпирических ортогональных вейвлетов и оценить их сходство для разных патологий. В-третьих, провести классификацию представления компонентов профиля пациента ортогональными вейвлетами на основе методов машинного обучения. В-четвертых, построить компоненты персонального профиля пациента в форме спектральных характеристик дыхательных шумов на основе методов машинного обучения.

ИСХОДНЫЕ ДАННЫЕ

В ходе проводимых исследований была использована база данных дыхательных шумов (Respiratory Sound Database, <https://www.kaggle.com/datasets/vbookshelf/respiratory-sound-database>) объемом 4 гигабайта. Она включает в себя 920 аннотированных записей продолжительностью от 10 до 90 сек. Эти записи были взяты у 126 пациентов. База данных содержит 5,5 час. записей, содержащих 6898 дыхательных циклов. Частота дискретизации равна 22050 Гц. Использовались пациенты из всех возрастных групп – дети, взрослые и пожилые люди [Roc18]. Основной объем выборки (80%) составили пациенты 60+.

СОВРЕМЕННОЕ СОСТОЯНИЕ

Автоматическая классификация дыхательных шумов включает в себя выделение и классификацию признаков [Ear00]. В качестве признаков чаще всего используются MFCC – коэффициенты, спектральные характеристики звуков, энергия и энтропия. Иногда используются коэффициенты вейвлетов Хаара и Добеши. Из методов классификации широкое распространение получили искусственные нейронные сети, метод опорных векторов, смешанные гауссовские модели, метод k-ближайших соседей и логистическая регрессия.

Источники возникновения дыхательных шумов являются одним из наиболее важных феноменов, определяющих диагностическую значимость диагностических исследований в пульмонологии. В настоящее время имеется ряд публикаций в которых рассматриваются физические причины возникновения дыхательных шумов.

Одной из первых публикаций на эту тему является работа Гротберга [Gro89]. В данной статье представлена модель образования везикулярных шумов в легких на основе динамического флаттера (рис. 1). При этом формируется задача идентификации параметров двухкомпонентной модели на основе получаемых данных измерений потоков воздуха на вдохе и на выдохе. Представлен математический анализ потока через гибкий канал, который рассматривается как модель вызванных потоком флаттерных колебаний, и вызывает свистящие звуки дыхания. Модель обеспечивает предсказание критической скорости потока, которая инициирует волны флаттера стенки, а также их частоты и длины волны. Математические результаты разделены на линейную теорию (малые колебания) и нелинейную теорию (большие колебания). Линейная теория определяет начало флаттера, в то время как нелинейная теория определяет взаимосвязь между скоростью потока и амплитудами, и частотами волн. Предсказания линейной теории хорошо коррелируют с данными, полученными в начале флаттера и ограничения потока во время экспериментов с воздушным потоком в толстостенных складных трубах. Предсказания нелинейной теории хорошо коррелируют с данными, полученными по мере того, как эти потоки разгоняются до более высоких скоростей при сохранении постоянной скорости потока. Конкретные диапазоны параметров выбираются для исследования и обсуждения применения к потокам в дыхательных путях. Согласно этой теории, механизм образования хрипов основан на взаимодействии сил течения и трения, а также сил упругого восстановления стенки и демпфирования. В частности, объясняется необходимость фазовой задержки между давлением потока и движением стенки.

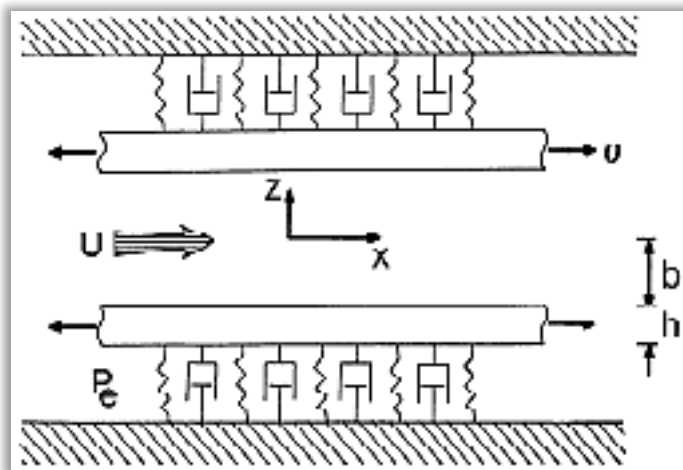


Рис. 1 Модель образования везикулярных шумов в легких на основе формирования самовозбуждающихся колебаний в гибком воздухопроводе [Gro89]

В известной отечественной работе [Кор97] представлен физический подход к описанию процесса колебания легких и возможных причин возникновения дыхательных шумов, который развит в работе [Кор04]. Модель, предложенная в работе [Кор04], предполагает формирование нескольких акустических резонансных контуров, автоколебательные режимы которых формируют свою часть спектра шумов. На рис. 2 представлен типичный спектр шума выдоха в соответствии с моделью. В предположении о взаимной маскировке акустических сигналов воздушного (по просвету дыхательных путей) и структурного (по тканям) проведения предложена акустическая модель распространения дыхательных звуков в респираторном тракте. Разработан оригинальный метод акустической интенсиметрии, который позволил впервые разделить спектральные составляющие воздушного и структурного проведения голоса и дыхательных шумов на грудную стенку. Существенно уточнена акустическая картина шумообразования дыхательных звуков. В частности, выяснено происхождение везикулярных звуков и свистов, в том числе при форсированном выдохе.

Данное направление развито в ряде работ под руководством В. И. Коренбаума, например, [Кор05], в которых исследуются акустические модели разных отделов дыхательной системы как источники формирования дыхательных шумов.

Фундаментальный обзор [Дья12] посвящен обсуждению публикаций, в которых представлены разные механизмы возникновения практически всех феноменов возникновения дыхательных шумов. Отмечается, что основными источниками шумов являются как естественные анатомо-физиологические особенности строения бронхолегочной системы, так и функциональные и анатомические изменения, вызванные разными синдромами и заболеваниями.

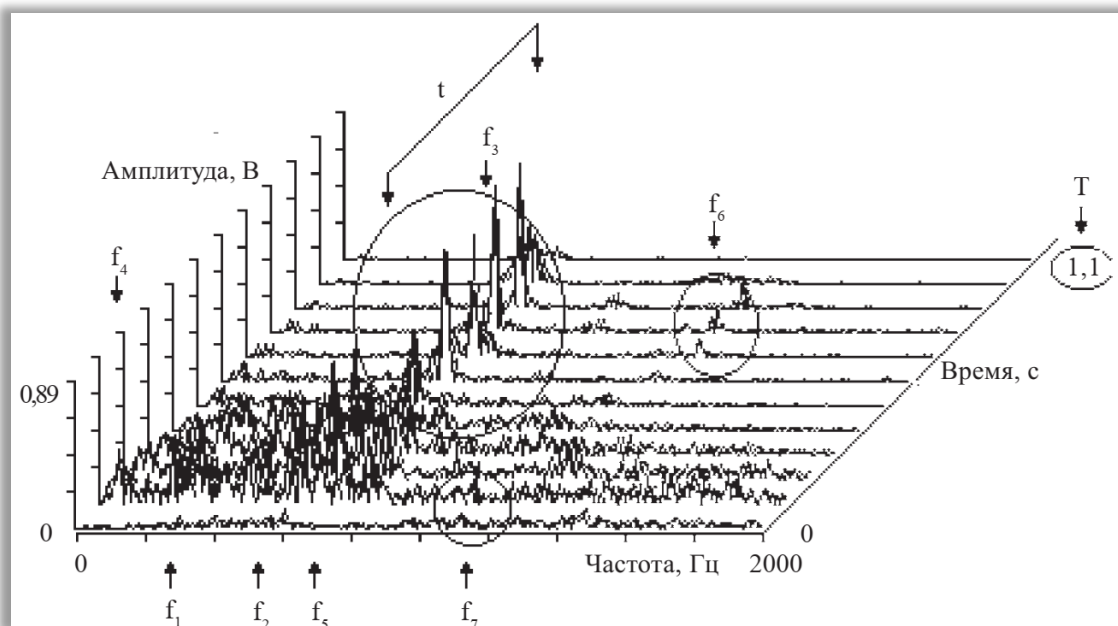


Рис. 2 Типичный рельефный спектр шума форсированного выдоха, зарегистрированный на трахее (эллипсами выделены диагностически значимые параметры) [Кор04]:

- f_1 — шум турбулентного потока в трахее; f_2 — шум турбулентного потока в главных бронхах;
- f_3 — свист форсированного выдоха (предположительно срыв вихрей) в диапазоне 400–600 Гц;
- t — продолжительность свиста форсированного выдоха в диапазоне частот 400–600 Гц;
- f_4 — низкочастотный свист автоколебательного происхождения;
- f_5 — среднечастотный свист автоколебательного происхождения;
- f_6 — высокочастотный свист автоколебательного происхождения в конце форсированного выдоха;
- f_7 — высокочастотный свист автоколебательного происхождения в начале форсированного выдоха.

Однако во всех этих работах недостаточно исследованы современные методы обработки данных, позволяющие провести анализ и обеспечить формирования диагностических значимых признаков необходимых во врачебной деятельности при меньших затратах времени. В связи с этим представляет интерес исследование имеющегося массива результатов респираторных исследований разных научных групп, содержащих собственно кривые дыхания и включающие дыхательные шумы различного происхождения для получения значений параметров дыхательной системы, имеющие значимость как для исследователей, так для практических специалистов.

Также следует отметить, что в ходе исследования авторы опирались на работы [Нас22, Саз23, Шап23, Юсу23, Зин24, Кор24, Мир24].

МЕТОДИКА ИССЛЕДОВАНИЯ

Было исследовано три алгоритма распознавания дыхательных шумов с целью построения персональных профилей пациентов с бронхолегочными заболеваниями. Два алгоритма были связаны с ортогональными вейвлетами, один алгоритм был основан на спектральных характеристиках звуков.

Для проведения вычислительных экспериментов было разработано программное обеспечение на языке Python. С его помощью осуществлялась предварительная обработка дыхательных шумов, выделения признаков и их классификация. Был разработан алгоритм балансировки классов. Для классификации признаков, выделенных из дыхательных шумов использовались метод k -ближайших соседей, наивный байесовский, дерево решений, случайный лес, градиентный бустинг, AdaBoost, XG Boost, Extra Boost, метод опорных векторов. Была использована вычислительная система в составе: Intel Core i9, 256 Gb RAM, NVIDIA RTX 3080 Ti.

Построение компонентов персонального профиля пациента на основе частотного анализа дыхательных шумов

Один из наиболее развитых подходов представлен в работах [Кор04, Кор05]. В соответствии с ним выделяются несколько источников шума в дыхательной системе, зависящих от скорости потока. Например, при малых скоростях $U/d \ll 200 \text{ с}^{-1}$ (U – линейная скорость, d – диаметр трубы) большей правдоподобностью обладает гипотеза о шуме турбулентного пограничного слоя со спектром, определяемым соотношением

$$\frac{2\pi\delta f}{8U} \leq 1, \quad (1)$$

где δ – толщина пограничного слоя, выше которого наблюдается резкий спад; U – линейная скорость потока; f – частота. Частота f будет определяться как

$$f \leq \frac{8U}{2\pi\delta}, \quad (2)$$

Соответственно, при скоростях $U/d > 200 \text{ с}^{-1}$ для нахождения спектрального максимума в качестве условия используется гипотеза о шуме турбулентной струи:

$$\frac{fd}{U} \approx 0,2. \quad (3)$$

Для второго случая выражение для спектрального максимума определяется как

$$f \approx \frac{0,2V}{N_i d_i^3}, \quad (4)$$

где N_i – количество бронхов i -й итерации, d_i – диаметр бронха i -й итерации. Соответственно, характерный диаметр бронха i -й итерации

$$d_i \approx (0,2V/N_i f)^3. \quad (5)$$

Основное предположение исследования заключается в том, что определение диаметра бронха, являющегося основным источником шума, позволяет оценить «глубину» воспалительного процесса в иерархии бронхолегочного дерева. Это позволяет определить характерные изменения в пульмонологическом профиле пациента. Такие изменения могут обеспечить повышение качества диагностики.

Поскольку характерные изменения связаны с максимум частотного спектра, то повышения качества диагностики достаточно определить этот максимум спектра. Сам спектр может быть получен типовыми методами частотного анализа, например, с помощью преобразования Фурье.

В работе в качестве групп сравнения использовались данные следующих: пациенты с астмой, с бронхоэктазией, ХОБЛ, с воспалением нижних дыхательных путей, с пневмонией, воспалением верхних дыхательных путей. Для данных этих групп рассчитывался спектр и определялась максимальная амплитуда.

Типовые спектры приведены на рис. 3.

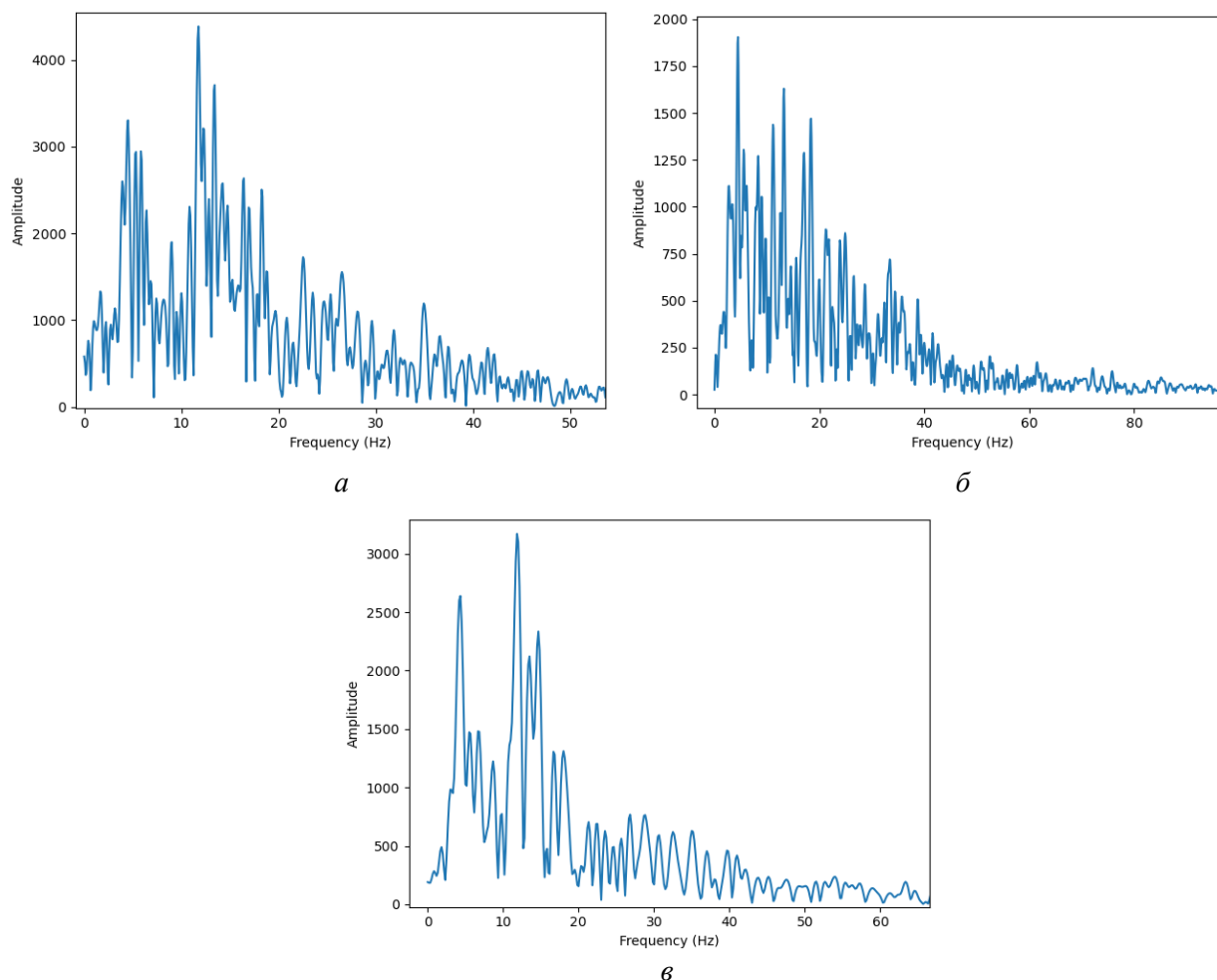


Рис. 3 Типовой спектр Фурье для разных видов заболеваний:
а – ХОБЛ; *б* – пневмония; *в* – здоровый (контроль)

Для полученных спектров определены примерные значения максимумов амплитуд частот: для астмы – 12,3 Гц; бронхоэктазии – 4,6 Гц; ХОБЛ – 11,6 Гц; пневмонии – 5,2 Гц; ВНДП – 16,2 Гц; ВВДП – 12,6 Гц; здоровый – 11,3 Гц. При таких значениях найти различия между некоторыми группами представляется нетривиальной задачей.

В качестве дополнительного инструмента для классификации видов заболеваний можно было бы использовать сами спектры частот. Однако из графиков видно, что различаются как амплитуды и площадь под кривыми. Такой способ может привести к затруднениям масштаба при анализе результатов множества измерений при слабом контроле уровня воспроизведения методики.

Таким образом, требуется применение инструмента анализа, максимально «унифицированного» по отношению к измеряемым данным. В качестве такового предлагается использовать подход, основанный на эмпирических ортогональных вейвлетах.

Построение компонентов персонального профиля пациента в форме эмпирических ортогональных вейвлетов и оценка сходства на основе корреляционной функции

Разработке теоретических основ вейвлетов и приложениям вейвлет-анализа посвящено большое количество работ, например, [Доб01, Вор99, Мал89].

Вместе с тем, необходимость анализа сигналов сложной природы, порождающих, в свою очередь, сигналы сложной формы, требует применения подходов, отличных от классического понимания вейвлетов. Наиболее эффективным с практической точки зрения является направления исследования, основанные на применении эмпирических вейвлетов. Расширение такого

подхода основано на сохранении аксиоматики вейвлет-представления сигналов и их вейвлет-анализа [Gil13, Има15].

Дальнейшее расширение практического применения вейвлет-анализа возможно в сторону использования таблично-заданных функций, часто не имеющих аналитического представления. Таким образом, вейвлет можно рассматривать, как конечную таблично-заданную функцию, обладающих свойствами, предъявляемыми к вейвлетам и определенным образом характеризующую сигнал. При этом возможно построение таблично-заданных ортогональных вейвлетов. Особенностью таких вейвлетов является то, что у них при заданных параметрах коэффициент корреляции с целевым сигналом стремится к единице, в то время как по отношению к другим сигналам корреляция этого вейвлета стремится к нулю.

При создании вейвлета выделяют характерную область исходного сигнала и устраняют избыточность путем детектирования сигнала. Размер рамки детектирования определяет степень разреженности исходного сигнала. Выделенная область детектированного сигнала нормируется, формируя вейвлет. Нормирование позволяет устранить влияние амплитуды сигнала на процесс распознавания.

Ортогональные вейвлеты можно использовать для распознавания неизвестных сигналов подобно фильтрам. В рамках проводимых исследований было установлено, что оптимальным размером вейвлета является 50 элементов, оптимальный размер рамки детектирования – 100 элементов. Рассмотрим процедуру создания вейвлета на примере везикулярного дыхательного шума. Открываем исходный звуковой сигнал и записываем его в массив A (рис 4).

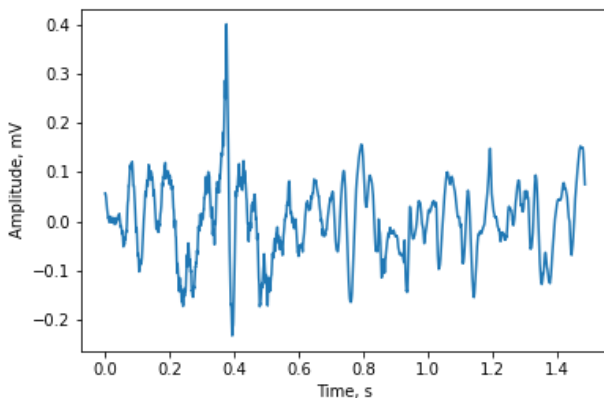


Рис. 4 Образец везикулярного дыхательного шума. Частота дискретизации 22050 Гц

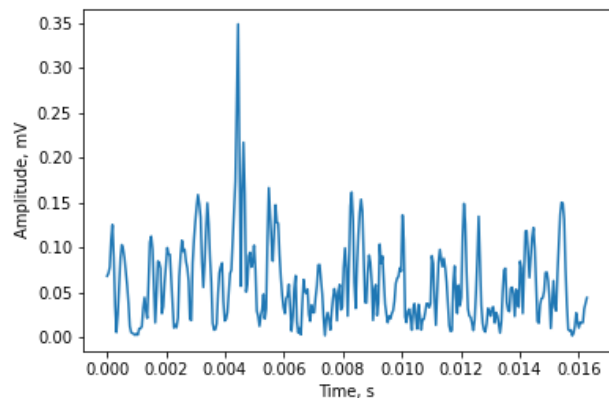


Рис. 5 Детектированный сигнал везикулярного дыхательного шума

Далее, осуществляли процедуру детектирования вейвлетов. Ниже показан детектированный сигнал везикулярного шума (рис. 5). Из рисунка видно, что размер детектированного сигнала в N раз меньше размера исходного сигнала.

После детектирование сигнала, определяем область, специфичную для конкретного сигнала. Была выбрана длине вейвлета в 50 элементов. Нормируя выделенная область, получаем вейвлет, представленный на рис. 6. Такую зависимость можно рассматривать как один из существенных компонентов профиля БЛЗ.

В процессе распознавания на вход системы подается неизвестный сигнал F . Далее производится детектирование сигнала так же, как и при создании вейвлета. Получаем массив детектированного сигнала C . Далее выбирается шаг распознавания, равный размеру вейвлета l . Внутри шага распознавания определяется коэффициент корреляции участка исследуемого сигнала C , равного l , со сравнимым вейвлетом. Найдя коэффициент корреляции на выбранном шаге, сдвигаем рамку распознавания на один элемент и повторяем процесс до исчерпания входного набора данных. Результат преобразования представлен на рис. 7.

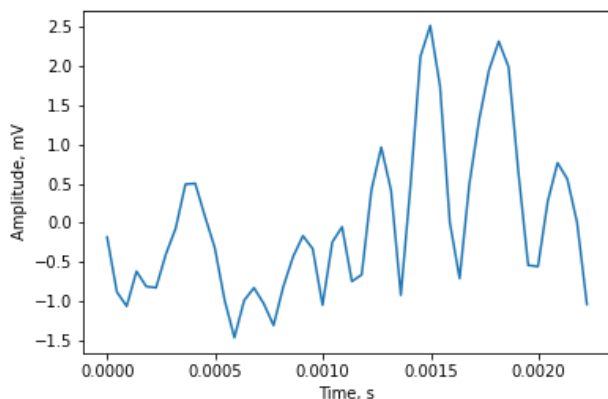


Рис. 6 Вейвлет везикулярного дыхательного шума

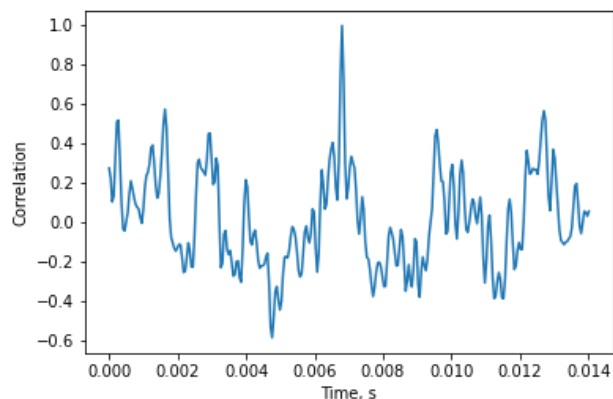


Рис. 7 Корреляционная функция вейвлета здорового испытуемого к целевому сигналу

В качестве меры сходства использован коэффициент корреляции базисной функций вейвлетов тестового набора данных и каждого участка исследуемого сигнала. Если коэффициент корреляции неизвестного сигнала и известного вейвлета близок к единице, то программа принимает решение об успехе распознавания. На рис. 6 показана корреляционная функция вейвлета везикулярного дыхания здорового испытуемого к целевому сигналу. Видно, что на одном из участков корреляционная функция равна единице. Остальным участкам сигнала соответствуют гораздо меньшие значения коэффициентов корреляции. Это означает, что исследуемый сигнал принадлежит здоровому испытуемому.

На рис. 8 показана корреляционная базисной функция вейвлета здорового испытуемого к сигналу пациента с ХОБЛ. Получился ожидаемо относительно низкий уровень корреляции.

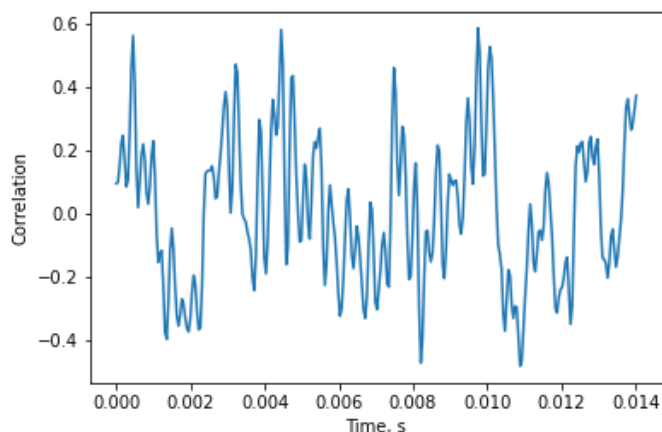


Рис. 8 Корреляционная функция вейвлета здорового испытуемого к сигналу пациента с хронической обструктивной болезнью легких. Максимальное значение 0,58

Построение компонента персонального профиля пациента в форме спектральных характеристиках дыхательных шумов с использованием методов машинного обучения.

Известно [Има15, Бог13, Дья03], что при решении задач в области защиты авторских прав на музыкальные произведения и при разработке программ классификации музыкальных произведений по жанрам в качестве признаков используются спектральные характеристики звуков. Используя прямую аналогию физической природы данных, проведена оценка эффективности применения подобного подхода при решении задачи классификации дыхательных шумов.

Для анализа использовались основные спектральные характеристики дыхательных шумов такие как среднеквадратичную энергию (gmse), спектрограмму мощности, спектральный центроид, спектральную полосу пропускания p -го порядка, частоту спада, скорость пересечения нуля временного аудио ряда, мел-частотные кепстральные коэффициенты (MFCC) (всего 26 признаков) [Бог13].

Проиллюстрируем упрощенный вариант классификации только на основе метрики Ассигасу, то есть доли правильных попаданий, которая приведена в табл. 1.

Из результатов, представленных в таблице видно, что наилучшей обобщающей способностью обладают такие методы классификации, как случайный лес и экстрасуст (наибольшая сумма значений по столбцам) и могут быть рекомендованы к использованию при классификации. Самыми трудными для диагностики диагнозами оказались хроническая обструктивная болезнь легких и воспаление верхних дыхательных путей (наименьшая сумма значений по строкам).

Таблица 1

Диагноз	knn	Random Forest	Ada Boost	Gradient Boosting	XG Boost	Extra Boost	SVM
Астма	1,00	1,00	1,00	1,00	0,99	1,00	0,88
Бронхоэктазия	1,00	1,00	0,99	1,00	0,96	1,00	0,78
ХОБЛ	0,86	0,93	0,82	0,89	0,81	0,94	0,70
Здоровый	1,00	1,00	0,92	0,98	0,89	1,00	0,73
ВНДП	1,00	1,00	1,00	1,00	0,99	1,00	0,81
Пневмония	1,00	1,00	0,95	0,96	0,94	1,00	0,81
ВВДП	0,80	0,92	0,86	0,89	0,86	0,92	0,74
Сумма	6,66	6,85	6,54	6,72	6,44	6,85	5,45

ЗАКЛЮЧЕНИЕ

Дыхательные шумы являются важным источником информации, позволяющим судить о состоянии органов дыхания. Появление на рынке цифровых стетоскопов делает возможным разработку телемедицинских сервисов онлайн-диагностики органов дыхания. Были проведены исследования в области разработки компонентов персональных профилей пациентов с бронхолегочными заболеваниями. Было использовано три подхода анализа дыхательных шумов, два из которых были основаны на ортогональных вейвлетах, а один — на спектральных характеристиках дыхательных шумов. Было установлено, что ортогональные вейвлеты обладают высокой селективностью к различным физиологическим состояниям органов дыхания и могут использоваться как важный компонент персонального профиля. Оценка эффективности классификации бронхолегочных заболеваний, когда в качестве признаков использовались вейвлеты показала, что высокой обобщающей способностью обладают такие методы, как случайный лес и экстрасуст. Аналогичный результат был получен при использовании в качестве признаков таких спектральных характеристик дыхательных шумов как среднеквадратичная энергия (gmse), спектрограмма мощности, спектральный центроид, спектральная полоса пропускания p -го порядка, частота спада, скорость пересечения нуля временного аудио ряда, мел-частотные кепстральные коэффициенты (MFCC). Предложенные подходы делают возможным осуществление скрининга пациентов с бронхолегочными заболеваниями, формирования их персональных профилей и отслеживания изменения таких профилей в динамике.

БЛАГОДАРНОСТИ И ПОДДЕРЖКА

Работа выполнена при финансовой поддержке РФФИ, грант 22-19-00471.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- [Cha16] Chamberlain D., Kodgule R., et al. "Application of semi-supervised deep learning to lung sound analysis" // 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2016;804-7. DOI: [10.1109/EMBC.2016.7590823](https://doi.org/10.1109/EMBC.2016.7590823).
- [Cre25] Creating a Patient Profile: Why It's Important for Hospitals to Understand their Patients. URL: <https://www.excel-medical.com/creating-a-patient-profile-why-its-important-for-hospitals-to-understand-their-patients>.
- [Ear00] Earis J, Cheetham B. "Current methods used for computerized respiratory sound analysis" // Eur Respir Rev. 2000; 10(77):586-90.
- [Efr79] Bradley Efron. "Bootstrap methods: Another look at the jackknife" // Annals of Statistics. 1979. Vol. 7, no. 1, pp. 1-26. DOI: [10.1214/aos/1176344552](https://doi.org/10.1214/aos/1176344552).
- [Gil13] Gilles J. "Empirical wavelet transform" // IEEE Transactions on Signal Processing. 2013. Vol. 61, no. 16, pp. 3999-4010. DOI: [10.1109/TSP.2013.2265222](https://doi.org/10.1109/TSP.2013.2265222).
- [Gro89] Grotberg J., Gavriely N. (1989). "Flutter in collapsible tubes: A theoretical model of wheezes" // Journal of Applied Physiology (Bethesda, Md.: 1985). 66. 2262-73. DOI: [10.1152/jappl.1989.66.5.2262](https://doi.org/10.1152/jappl.1989.66.5.2262).
- [Mal89] Mallat S. "A theory for multiresolutional signal decomposition: the wavelet representation" // IEEE Trans. Pattern Analysis and Machine Intelligence, 1989, no. 7, pp. 674-693. DOI: [10.1109/34.192463](https://doi.org/10.1109/34.192463).
- [Mar14] Marques A, Oliveira A, Jácome C. "Computerized adventitious respiratory sounds as outcome measures for respiratory therapy: a systematic review" // Respir Care. 2014 May;59(5):765-76. DOI: [10.4187/respcare.02765](https://doi.org/10.4187/respcare.02765).
- [Pii95] Piirila P., Sovijarvi A. R. "Crackles: recording, analysis and clinical significance" // Eur Respir J. Eur Respiratory Soc; 1995;8(12):2139-48. DOI: [10.1183/09031936.95.08122139](https://doi.org/10.1183/09031936.95.08122139).
- [Pra17] Pramono R. X. A., Bowyer S., Rodriguez-Villegas E. "Automatic adventitious respiratory sound analysis: A systematic review" // PloS One. Public Library of Science; 2017; 12(5):e0177926. DOI: [10.1371/journal.pone.0177926](https://doi.org/10.1371/journal.pone.0177926).
- [Roc] B. M. Rocha et al. A Respiratory Sound Database for the Development of Automated Classification. URL: <https://bhchallenge.med.auth.gr>.
- [Roc18] Rocha B. M., Filos D., et al. (2018). "A respiratory sound database for the development of automated classification" // In Precision Medicine Powered by pHHealth and Connected Health, pp. 51-55. Springer, Singapore. DOI: [10.1007/978-981-10-7419-6_6](https://doi.org/10.1007/978-981-10-7419-6_6).
- [Sar15] Sarkar M., Madabhavi I., Niranjana N., Dogra M. "Auscultation of the respiratory system" // Ann Thorac Med. Medknow Publications; 2015;10(3):158. DOI: [10.4103/1817-1737.160831](https://doi.org/10.4103/1817-1737.160831).
- [Sar19] Murat Sariyar and Irene Schlünder. "Challenges and legal gaps of genetic profiling in the era of Big Data" // Front. Big Data, 12 November 2019. Sec. Big Data Networks. DOI: [10.3389/fdata.2019.00040](https://doi.org/10.3389/fdata.2019.00040).
- [Sch06] Leslie O. Schulz, Peter H. Bennett, et al. "Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the U.S." // Diabetes Care 1 August 2006; 29 (8): 1866-1871. DOI: [10.2337/dc06-0138](https://doi.org/10.2337/dc06-0138).
- [Smi88] Smith J. W., Everhart J. E., et al. (1988). "Using the ADAP learning algorithm to forecast the onset of diabetes mellitus" // In Proceedings of the Symposium on Computer Applications and Medical Care (pp. 261-265). IEEE Computer Society Press. URL: <https://www.researchgate.net/publication/248284447>.
- [Sov00] Sovijarvi A. R. A. et al. "Characteristics of breath sounds and adventitious respiratory sounds" // European Respiratory Review. 2000; 10(77):591-596.
- [WHO17] World Health Organization. The top 10 causes of death. 2017.
- [Бог13] Богданов М. Р., Думчиков А. А. Способ изучения разнообразия орнитофауны методом вейвлет-анализа // Известия Самарск. науч. центра РАН. 2013. Т. 15, № 3 (4). С. 1232-1235. EDN: [SAENFJ](https://www.edn.ru/SAENFJ). [[Bogdanov M. R., Dumchikov A. A. "Method of studying the diversity of avifauna by wavelet analysis" // Bulletin of the Samara Scientific Center of the Russian Academy of Sciences. 2013. Vol. 15, no. 3 (4), pp. 1232-1235. EDN: [SAENFJ](https://www.edn.ru/SAENFJ). (In Russian).]]
- [Вор99] Воробьев В. И., Грибунин В. Г. Теория и практика вейвлет-преобразования. СПб: ВУС, 1999. EDN: [RCNCFE](https://www.edn.ru/RCNCFE). [[Vorobiev V. I., Gribunin V. G. Theory and Practice of Wavelet Transform. St. Petersburg: VUS, 1999. EDN: [RCNCFE](https://www.edn.ru/RCNCFE). (In Russian).]]
- [Доб01] Добеши И. Десять лекций по вейвлетам. М.: РХД, 2001. [[Daubechies I. Ten Lectures on Wavelets. Moscow: RHD, 2001. (In Russian).]]
- [Дья03] Дьяченко А. И. Математические модели механики легких с распределенными параметрами: Автореф. дисс. ... д-ра техн. наук. М., 2003. [[Dyachenko A. I. Mathematical Models of Lung Mechanics with Distributed Parameters: Abstract of a Doctor of Technical Sciences dissertation. Moscow, 2003. (In Russian).]]
- [Дья12] Дьяченко А. И., Михайловская А. Н. Респираторная акустика (обзор) // Труды ИОФАН. 2012. С. 136-181. EDN: [PFPTRX](https://www.edn.ru/PFPTRX). [[Dyachenko A. I., Mikhailovskaya A. N. "Respiratory acoustics (review)" // Proc. General Physics Institute. 2012. Pp. 136-181. EDN: [PFPTRX](https://www.edn.ru/PFPTRX). (In Russian).]]
- [Зин24] Зиновьев М. С., Нургаянова О. С. Прогнозирование вероятности развития диабетической ретинопатии у пациентов с сахарным диабетом: анализ методов машинного обучения // СИИТ. 2024. Т. 6, № 3(18). С. 95-101. DOI [10.54708/2658-5014-SIIT-2024-no3-p95](https://doi.org/10.54708/2658-5014-SIIT-2024-no3-p95). EDN: [VLFPLP](https://www.edn.ru/VLFPLP).
- [Има15] Имамвердиев Я. Н., Сухостат Л. В. Метод оценки периода основного тона с применением эмпирического вейвлет преобразования // Radio Electronics, Computer Science, Control. 2015. № 2. DOI: [10.15588/1607-3274-2015-2-6](https://doi.org/10.15588/1607-3274-2015-2-6). [[Imamverdiev Ya. N., Sukhostat L. V. "Method for estimating the period of the fundamental tone using empirical wavelet

- transform" // Radio Electronics, Computer Science, Control. 2015. No. 2. DOI: [10.15588/1607-3274-2015-2-6](https://doi.org/10.15588/1607-3274-2015-2-6). (In Russian).]]
- [Кор04] Коренбаум В. И., Почекутова И. А. и др. Акустическая диагностика системы дыхания человека на основе объективного анализа дыхательных звуков // Вестник ДВО РАН. 2004. № 5. С. 68-79. EDN: [HPNCBL](#). [[Korenbaum V. I., Pochekutova I. A. et al. "Acoustic diagnostics of the human respiratory system based on objective analysis of respiratory sounds" // Bulletin of the Far Eastern Branch of the Russian Academy of Sciences. 2004. No. 5, pp. 68-79. EDN: [HPNCBL](#). (In Russian).]]
- [Кор05] Коренбаум В. И., Почекутова И. А. Акустико-биомеханические взаимосвязи в формировании шумов форсированного выдоха человека. Владивосток: Дальнаука, 2005. 150 с. [[Korenbaum V. I., Pochekutova I. A. Acoustic-Biomechanical Relationships in the Formation of Human Forced Expiratory Noise. Vladivostok: Dalnauka, 2005. (In Russian).]]
- [Кор24] Коровин Е. А., Чиглицева С. А., Сазонова Е. Ю., Сметанина О. Н. Медицинская рекомендательная система на основе автоматического извлечения знаний из текстов // СИИТ. 2024. Т. 6, № 4(19). С. 111-121. DOI [10.54708/2658-5014-SIIT-2024-no4-p111](https://doi.org/10.54708/2658-5014-SIIT-2024-no4-p111). EDN: [OTVTXR](#).
- [Кор97] Коренбаум В. И., Кулаков Ю. В., Тагильцев А. А. Акустические эффекты в системе дыхания человека при форсированном выдохе // Акустический журнал. 1997. Т. 43, № 1. С. 78-86. EDN: [UQWXTT](#). [[Korenbaum V. I., Kulakov Yu. V., Tagiltsev A. A. "Acoustic effects in the human respiratory system during forced exhalation" // Acoustic Journal. 1997. Vol. 43, no. 1, pp. 78-86. EDN: [UQWXTT](#). (In Russian).]]
- [Мир24] Мирасов О. О., Шахмаметова Г. Р. Сетевой анализ профилей экспрессии генов // СИИТ. 2024. Т. 6, № 3(18). С. 61-70. EDN: [RCXENH](#). [[Mirasov O. O., Shakhmametova G. R. "Network analysis of gene expression profiles" // SIIT. 2024. Vol. 6, no. 3(18), pp. 61-70. EDN: [RCXENH](#). (In Russian).]]
- [Нас22] Насыров Р. В. Причинный подход к построению бионических вычислений на основе рекурсивных моделей анализа данных // СИИТ. 2022. Т. 4, № 1(8). С. 27-36. EDN: [UOMMOU](#). [[Nasyrov R. V. "Causal approach to the construction of bionic computing based on recursive models of data analysis" // SIIT. 2022. Vol. 4, no. 1(8), pp. 27-36. EDN: [UOMMOU](#). (In Russian).]]
- [Саз23] Сазонова Е. Ю., Сметанина О. Н., Журавлева К. И., Юлаев Р. С. Интеллектуальная СППР при управлении психофизическим состоянием человека // СИИТ. 2023. Т. 5, № 6(15). С. 38-49. DOI [10.54708/2658-5014-SIIT-2023-no5-p38](https://doi.org/10.54708/2658-5014-SIIT-2023-no5-p38). EDN: [JFLRCX](#).
- [Шап23] Шапошникова А. С., Богданов М. Р. Определение сердечного ритма плода по неинвазивному ЭКГ с применением различных фильтров // СИИТ. 2023. Т. 5, № 6(15). С. 32-37. EDN: [WBBOVK](#). [[Shaposhnikova A. S., Bogdanov M. R. "Determination of fetal heart rate by non-invasive ECG using various filters" // SIIT. 2023. Vol. 5, no. 6(15), pp. 32-37. EDN: [WBBOVK](#). (In Russian).]]
- [Юсу23] Юсупова Н. И., Нургаянова О. С., Зулкарнеев Р. Х. Формализация этапов риск-анализа в СППР с учетом оценок клинических рисков при бронхолегочных заболеваниях // СИИТ. 2023. Т. 5, № 1(10). С. 11-24. DOI [10.54708/2658-5014-SIIT-2023-no1-p11](https://doi.org/10.54708/2658-5014-SIIT-2023-no1-p11). EDN: [KHIIHT](#).

МЕТАДАТА | МЕТАДАННЫЕ

Поступила в редакцию 28 апреля 2025 г.

Title: Personal profiles of patients with bronchopulmonary diseases.

Abstract: The research in the field of development of personal profiles of patients with bronchopulmonary diseases was conducted. To determine the profile components, three approaches to the analysis of respiratory noises were used, two of which were based on orthogonal wavelets, and one on the spectral characteristics of respiratory noises. It was found that orthogonal wavelets have high selectivity to various physiological states of the respiratory organs. Evaluation of the classification efficiency of bronchopulmonary diseases, when wavelets were used as features, showed that such methods as random forest and extra boost have high generalizing ability. A similar result was obtained when using such spectral characteristics of respiratory noises as features as root-mean-square energy (rmse), power spectrogram, spectral centroid, p-order spectral bandwidth, rolloff frequency, zero-crossing rate of the time audio series, small-frequency cepstral coefficients (MFCC). The proposed approaches make it possible to form components of a patient's personal profile based on screening of patients with bronchopulmonary diseases.

Key words: personal patient profile; bronchopulmonary diseases; machine learning; wavelet analysis.

Language: Russian.

Об авторах | About the authors

БОГДАНОВ Марат Робертович

Уфимский университет науки и технологий, Россия.
Доц. каф. вычислительной математики и кибернетики. Дипл. биолог, преподаватель биологии и химии (Баш. гос. ун-т, 1991). Канд. биол. наук (Всерос. НИИ защиты растений, 1997). Иссл. в обл. интелл. анализа биомедицинских данных.
E-mail: bogdanov_marat@mail.ru
ORCID: [0000-0003-1775-5920](https://orcid.org/0000-0003-1775-5920)

BOGDANOV Marat Robertovich

Ufa University of Science and Technologies, Russia.
Docent, Dept. of Geographic Information Systems. Dipl. biologist, teacher of biology and chemistry (Bashkir State Univ., 1991). Cand. Biol. Sci. (All-Russ. Inst. of Plant Protection, 1997). Research in the field of biomedical data mining.
E-mail: bogdanov_marat@mail.ru
ORCID: [0000-0003-1775-5920](https://orcid.org/0000-0003-1775-5920)

НАСЫРОВ Рашит Вильевич

Уфимский университет науки и технологий, Россия.

Доц. каф. технической кибернетики. Дипл. инж.-системотехник (Уфимск. авиац. ин-т, 1989). Канд. техн. наук по автоматиз. системам принятия решений (Уфимск. гос. авиац. техн. ун-т, 1994). Канд. фарм. наук по моделям и технологиям лекарств. обеспечения в чрезв. ситуациях (РУДН, 2005). Иссл. в обл. обработки эксперим. и стат. данных, функц. биомеханики, моделей и алгоритмов принятия решений.

E-mail: nrash@yandex.ru

ШАХМАМЕТОВА Гюзель Радиковна

Уфимский университет науки и технологий, Россия.

Зав. каф. вычислительной математики и кибернетики. Дипл. инж. по инф. системам (Уфимск. авиац. ин-т, 1992). Д-р техн. наук по сист. анализу, управлению и обработке информации (Уфимск. гос. авиац. техн. ун-т, 2013). Иссл. в обл. интеллект. поддержки принятия решений, распознавания образов, обработки биомед. данных методами машинного обучения и искусственного интеллекта.

E-mail: shakhgouzel@mail.ru

ORCID: <http://orcid.org/0000-0002-7742-793X>

NASYROV Rashit Vilyevich

Ufa University of Science and Technologies, Russia.

Assoc. Prof. of the Department of Technical Cybernetics. Dipl. Systems Engineer (UAI, 1989). Cand. of Techn. Sci. (Ufa State Aviation Techn. Univ., 1994). Cand. of Pharm. Sci. (RUDN Univ., 2005). Research in the field of experimental and statistical data processing, functional biomechanics, simulation and math modeling of systems and processes, models and algorithms for decision-making.

E-mail: nrash@yandex.ru

SHAKHMAMETOVA Gyuzel Radikovna

Ufa University of Science and Technologies, Russia.

Head of the Department of Computational Mathematics and Cybernetics. Dipl. Eng. on information systems (Ufa Aviation Inst., 1992). Dr. Tech. Sciences in system analysis, management and information processing (Ufa State Aviation Technical University, 2013). Research in the field of intelligent decision support, pattern recognition, processing of biomedical data using machine learning and artificial intelligence methods.

E-mail: shakhgouzel@mail.ru

ORCID: <http://orcid.org/0000-0002-7742-793X>