

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК 004.457

Родько  
Владислав Мирославович

Программный комплекс для автоматизации  
медицинской реабилитации пациентов

**АВТОРЕФЕРАТ**

на соискание академической степени  
магистра технических наук

по специальности 1-40 80 04 – Математическое моделирование, численные  
методы и комплексы программ

Научный руководитель  
Сиротко С.И.  
к.ф.-м.н., доцент

Минск 2019

## КРАТКОЕ ВВЕДЕНИЕ

Современная медицина развивается исключительно быстрыми шагами. На такое развитие, конечно, повлиял и огромный прогресс в сфере информационных технологий. Сегодня медицинские учреждения проводят лечение пациентов с самыми сложными заболеваниями, используя современное оборудование. Сфера здравоохранения – это высокотехнологичная отрасль, где успешно развиваются трансплантология и травматология, пластическая хирургия и онкология, нейрохирургия, офтальмология, гинекология, стоматология и другие области, которые позволяют спасать жизни ранее безнадежных больных. Значительно улучшилась техническая оснащённость медицинских учреждений, появилась возможность диагностировать заболевание на самой ранней стадии и быстро восстанавливать работоспособность пациентов. Привычными стали малоинвазивные процедуры с использованием эндоскопической аппаратуры, микрохирургия и лазерная коррекция зрения, пересадка органов и тканей, исправление врожденных и приобретенных дефектов.

Но все же некоторые проблемы до сих пор остаются актуальными. Для их решения на помощь приходит Интернет-вещей. Интернет вещей (Internet of things, IoT) – это один из наиболее популярных передовых технологий за последнее время. Возможность внедрения практических решений благодаря концепции интернета вещей дает преимущества во множестве сфер человеческой деятельности.

Аналитики Gartner предсказывают, что к 2020 г. количество подключенных устройств составит 20 млрд. По прогнозам IHS Global Insight, к 2020 г. число подключенных устройств составит всего 18 млрд, стартовав с отметки в 8 млрд в 2014 г..

Одно из главных направлений в интернете вещей – развитие медицины и здравоохранения. Для этого направления уже есть особые термины – например, «Интернет медицинских вещей» или «Интернет вещей» в здравоохранении.

В системе здравоохранения особенно остро встает вопрос наблюдения за пациентами, отслеживание их местоположения и состояния, а так же слежение за самим медицинским учреждением, его внутренним микроклиматом. Именно в этой отрасли, благодаря развитию Интернета вещей, цифровое преобразование пойдет высокими темпами.

По прогнозам исследователей (компания Allied Market Research), рынок медицинских IoT-гаджетов и IoT-приложений до 2021 года вырастет до \$136,8 млрд. Среднегодовой темп роста рынка медицинского Интернета вещей составит 12,5%. Кроме этого, в ближайшем будущем в сфере здравоохранения произойдет рост количества высокотехнологичных сервисов и систем. Драйверами рынка станут: повышение доступности высокоскоростного Интернета, снижение цен на датчики, распространение медицинских гаджетов и рост информированности пользователей об инновациях в медицине.

Таким образом, традиционную модель «пациент-доктор» ожидают значительные изменения, а медицинские устройства открывают новую эру в цифровом здравоохранении.

Именно медицина и здравоохранение в целом становятся драйвером роста сегмента интернета вещей по одной главной причине – здесь возможно массовое и прямое взаимодействие человека и компактного электронного прибора. Востребованность этих технологий будет расти, поскольку интернет вещей позволяет существенно повысить качество и доступность медицинских услуг, снизить расходы на медицинское обслуживание для пациентов, даст им возможность получать консультации от лучших врачей и передовое медицинское обслуживание в тех регионах, где пока нет больниц с высокотехнологичным оборудованием.

Согласно исследованию консалтингового агентства McKinsey использование аналитики больших данных позволит системе здравоохранения сэкономить от 350 до 450 миллиардов долларов, оптимизируя расходы по использованию медицинского оборудования, рабочего времени врачей и назначения фармакологических препаратов.

Современные медицинские учреждения сталкиваются с широким спектром различных задач, например, необходимость постоянного набора персонала для повышения эффективности работы, наблюдение за физическими показателями пациентов. и др. И все эти проблемы имеют общую причину – отсутствие комплексного решения, которое будет выполнять постоянное и непрерывное измерение и контроль параметров среды, технологических процессов, статуса и состояния пациентов, персонала, медицинского и другого оборудования.

Современные устройства, подключённые к интернету вещей, сегодня становятся неотъемлемой частью жизни многих медицинских учреждений и значительно помогают в автоматизации процессов и позволяют сэкономить драгоценное время врачей. Использование сенсоров и удаленно работающих медицинских приборов предоставит пациентам доступ к передовым медицинским технологиям. Развитие удаленной диагностики позволит преодолевать географический разрыв в оказании медицинских услуг, повысит эффективность медицинской помощи в чрезвычайных ситуациях, а также это позволит значительно снизить стоимость оказываемых услуг. Объединенные с помощью Интернета медицинские устройства смогут собирать и передавать данные о пациентах на большом расстоянии и за считанные секунды, что значительно снизит количество медицинских ошибок.

# **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

## **Цель и задачи исследования**

*Целью* диссертационной работы является исследование и реализация программного комплекса для автоматизации медицинской реабилитации пациентов.

*Объектом* исследования являются процессы организации физической реабилитации.

*Предметом* исследования является программное обеспечение для автоматизации организации физической реабилитации.

## **Личный вклад соискателя**

Все изложенные в диссертации результаты исследования получены соискателем лично с учетом рекомендаций и замечаний научного руководителя.

## **Апробация результатов диссертации**

Основные положения диссертационной работы докладывались на 55-ой научной конференции аспирантов, магистрантов и студентов БГУИР (Минск, Беларусь, 2019).

## **Опубликованность результатов диссертации**

По теме диссертации опубликовано 3 печатные работы, из них 2 статьи в рецензируемом издании, 1 работа в сборниках трудов и материалов научных конференций.

## **Структура и объем диссертации**

Структура диссертационной работы обусловлена целью, задачами и логикой исследования. Состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников, списка публикаций автора и приложений.

Общий объем работы составляет 61 страницу, из которых основного текста – 42 страницы, включая 14 иллюстраций и 4 формулы, список использованных источников из 37 наименований на 3 страницах и 4 приложения на 12 страницах.

## ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

**В первой главе** проведен анализ предметной области, поиск и изучение подходов машинного обучения, которые могут быть использованы для анализа поведения пользователей сайтов и приложений. Выполнен анализ применяемых методов и алгоритмов обработки данных, выявлены их достоинства и недостатки. Итогом является выбор и описание выбранных подходов и алгоритмов.

Машинное обучение включают в себя множество различных типов методов, они отличаются своими достоинствами при применении к некоторым задачам и невозможность их применения к другим типам задач.

По способам решения задачи разделяют на обучение «с учителем» и «без учителя». Обучение «с учителем» требует существования ответов, с помощью которых алгоритм будет обучаться давать ответ. Обучение «без учителя» объединяет задачи, выявляющие описательные модели, например, закономерности в покупках, совершаемых клиентами большого магазина.

Для решения поставленных в диссертации для исследованием задач могут помочь методы обучения «без учителя», такие как кластеризация и анализ главных компонент.

**Вторая глава** посвящена разработке архитектуры программного обеспечения, подготовке тестовых данных для последующего анализа.

Для демонстрации работоспособности алгоритма анализа использован сгенерированный мной набор метрик поведения пользователей в системе управления проектами небольших групп. Использование подразумевает регистрацию, вход в аккаунт, открытие главной страницы, досок с карточками, создание карточек и т.д. При генерации этого набора были заданы два типа пользователей: корпоративные и персональные.

В процессе разработки системы были использованы следующие инструменты:

- Python, как язык разработки;
- Django, широко распространенный веб-фреймворк для разработки сайтов, веб-приложений, API на языке python;
- PostgreSQL в качестве сервера базы данных;
- Библиотека машинного обучения Scikit-learn;
- Библиотека Pandas;
- Библиотека Matplotlib.

С помощью вышеперечисленных инструментов удалось разработать алгоритм анализа поведения пользователей. Синтаксис Python является простым и лаконичным, имеет множество хороших и поддерживаемых библиотек, таких как scikit-learn, pandas, matplotlib.

В третьей главе описывается алгоритма решения поставленной задачи. Проведено исследование различных алгоритмов кластеризации, описаны подходы для улучшения качества результатов.

Алгоритм действий представлен на рисунке 1.



Рисунок 1 Алгоритм анализа поведения пользователей

При выполнении разработанного алгоритма происходят выборка информации из базы данных и её подготовка к дальнейшему анализу,

применение кластерного анализа, описание всех полученных групп, поиск изменений в группах в сравнении с предыдущей неделей, генерация отчета.

Первым шагом является подсчет метрик для каждого пользователя для анализа. Для этого нужно запросить все события пользователей из базы данных за последнюю неделю. Следующим шагом является агрегация выборки из предыдущего шага таким образом, чтобы теперь строки представляли собой информацию о событиях уникального пользователя. Здесь активно используется библиотека `pandas`.

Второй шаг представляет собой процесс разбиения пользователей на группы. Но для того чтобы выбрать наиболее подходящий алгоритм действий при кластеризации, было проведено исследование для выбора наиболее подходящего алгоритма кластерного анализа и метода по улучшению результатов кластеризации.

В процессе этого исследования было проведено:

1. сравнение трех алгоритмов кластеризации на качество результатов и быстродействие работы;
  2. исследование возможности улучшения результатов путем уменьшения размерности данных при помощи анализа главных компонент;
  3. исследование возможности применения алгоритма DBSCAN и фильтрации аномальных объектов для улучшения результатов кластеризации.
- Результаты исследования данного пункта представлены в [1-А, 2-А].

Следующим этапом является создание описания для каждой из групп – необходимо выделить наиболее значимые признаки группы. Это параметры, значения которых наиболее существенно отличаются от соответствующих значений других групп.

Далее происходит сравнение результатов анализа текущей недели с результатами прошлой. Этот шаг позволяет узнать изменения в пользовательских характеристиках, позволит заметить тенденции в изменении характеристик или числа пользователей различных групп.

# **ЗАКЛЮЧЕНИЕ**

## **Основные научные результаты диссертации**

1. Разработан алгоритм проведения анализа поведения пользователей с помощью МО. В ходе работы алгоритма происходит обработка действий пользователей за последнюю неделю, разделение пользователей на различные группы по поведенческим характеристикам, описание признаков полученных групп, поиск изменений в группах по сравнению с предыдущей неделей.

2. Разработан подход для улучшения эффективности кластеризации данных путем использованием анализа главных компонент (РСА). Предложен алгоритм сравнения работоспособности данного подхода в зависимости от количества выделяемых компонент.

3. Разработан подход для улучшения результатов кластерного анализа путем предварительного использования алгоритма кластеризации DBSCAN для фильтрации аномальных объектов в исходной выборке [1-А, 2-А].

## **Рекомендации по практическому использованию результатов**

1. Полученные результаты формируют теоретическую и практическую базу для разработки ПО для решения задачи анализа поведения пользователей, а точнее, поиска групп пользователей с помощью кластерного анализа, описания найденных групп, нахождения изменения поведения различных групп пользователей. Это позволяет их использовать в автоматическом режиме в системах веб-аналитики, позволяет уменьшить время проведение анализа человеком, отказаться от работы с графиками различных типов и данных.

2. Разработанные методы и алгоритмы анализа поведения пользователей могут применяться в автоматизированных системах веб-аналитики.



## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

[1–А.] Ковалёв С.П., Калугина М.А. Комбинация алгоритмов К-средних и DBSCAN при анализе поведения пользователей // Студенческий: электрон. научн. журн. 2019. № 18(62). URL: <https://sibac.info/journal/student/62/140295> (дата обращения: 15.05.2019).

[2–А.] Ковалёв С. П. Использование алгоритма кластеризации DBSCAN для фильтрации выбросов в данных / Ковалёв С. П. // 55-я юбилейная научная конференция аспирантов, магистрантов и студентов. - 2019. - № 4(20). - с. 198-200.