



# BIG DATA And Advanced Analytics 2023

2

**IX**

международная  
научно-практическая  
конференция



Министерство образования Республики Беларусь  
Учреждение образования «Белорусский государственный университет  
информатики и радиоэлектроники»  
Национальная академия наук Беларуси  
Государственное научное учреждение  
«Объединенный институт проблем информатики  
Национальной академии наук Беларуси»  
Ташкентский университет информационных технологий, Узбекистан  
Южный федеральный университет, Российская Федерация  
Северо-Кавказский федеральный университет, Российская Федерация  
Евразийский национальный университет имени Л. Н. Гумилева, Республика Казахстан  
Технологический университет Шарифа, Иран

## **BIG DATA И АНАЛИЗ ВЫСОКОГО УРОВНЯ**

## **BIG DATA AND ADVANCED ANALYTICS**

СБОРНИК НАУЧНЫХ СТАТЕЙ  
IX МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ

В двух частях

Часть 2

(Республика Беларусь, Минск, 17–18 мая 2023 года)

Минск БГУИР 2023



УДК 004.6(082)

ББК 32.973.3

Б59

#### **Редакционная коллегия:**

**В.А. Богуш** – доктор физико-математических наук, профессор, ректор учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», Республика Беларусь;

**С.К. Дик** – кандидат физико-математических наук, доцент, депутат Палаты представителей Национального собрания Республики Беларусь седьмого созыва, Республика Беларусь;

**Д.В. Лихачевский** – кандидат технических наук, доцент, декан факультета компьютерного проектирования учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», Республика Беларусь;

**Т.В. Казак** – доктор психологических наук, профессор, заведующий кафедрой инженерной психологии и эргономики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», Республика Беларусь

#### **Рецензенты:**

**А.В. Тузиков** – доктор физико-математических наук, профессор, член-корреспондент Национальной академии наук Беларуси, генеральный директор государственного научного учреждения «Объединенный институт проблем информатики Национальной академии наук Беларуси», Республика Беларусь;

**Д.А. Тусупов** – заведующий кафедрой информационных систем Евразийского национального университета имени Л. Н. Гумилева, доктор физико-математических наук, профессор, Республика Казахстан;

**Д.В. Лихачевский** – кандидат технических наук, доцент, декан факультета компьютерного проектирования учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», Республика Беларусь

**BIG DATA** и анализ высокого уровня = **BIG DATA and Advanced Analytics** : сб. науч. Б59 ст. IX Междунар. науч.-практ. конф. В 2 ч. Ч. 2 (Республика Беларусь, Минск, 17–18 мая 2023 года) / редкол. : В. А. Богуш [и др.]. – Минск : БГУИР, 2023. – 384 с.

ISBN 978-985-543-713-1 (ч. 2).

В сборнике опубликованы результаты научных исследований и разработок в области **BIG DATA and Advanced Analytics** для оптимизации ИТ- и бизнес-решений, а также тематических исследований в области медицины, образования и экологии.

УДК 004.6(082)

ББК 32.973.3

ISBN 978-985-543-713-1 (ч. 2)

ISBN 978-985-543-711-7

© УО «Белорусский государственный университет информатики и радиоэлектроники», 2023



## ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ



**Председатель, Богуш В.А.**

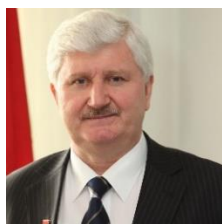
Ректор Белорусского государственного университета информатики и радиоэлектроники, доктор физико-математических наук, профессор



**Заместитель председателя, Дик С.К.**

Депутат Палаты представителей Национального собрания Республики Беларусь седьмого созыва, кандидат физико-математических наук, доцент, Республика Беларусь

## ЧЛЕНЫ ОРГАНИЗАЦИОННОГО КОМИТЕТА



**Батура М.П.**, научный руководитель НИЛ 8.1 БГУИР, доктор технических наук, профессор, академик «Международной академии наук высшей школы», заслуженный работник образования Республики Беларусь, Республика Беларусь



**Болдырев А.С.**, директор института радиотехнических систем и управления Южного федерального университета (г. Таганрог), кандидат физико-математических наук, доцент, Российская Федерация



**Давыдов М.В.**, первый проректор БГУИР, кандидат технических наук, доцент, Республика Беларусь



**Казак Т.В.**, заведующий кафедрой инженерной психологии и эргономики БГУИР, доктор психологических наук, профессор, Республика Беларусь





**Лихачевский Д.В.**, декан факультета компьютерного проектирования Белорусского государственного университета информатики и радиоэлектроники, кандидат технических наук, доцент, Республика Беларусь



**Пархименко В.А.**, заведующий кафедрой экономики БГУИР, кандидат экономических наук, доцент, Республика Беларусь



**Пискун Г.А.**, заместитель декана факультета компьютерного проектирования БГУИР, доцент кафедры проектирования информационно-компьютерных систем БГУИР, кандидат технических наук, доцент, Республика Беларусь



**Тузиков А.В.**, главный научный сотрудник Объединенного института проблем информатики Национальной академии наук Беларуси и доктор физико-математических наук Беларуси, профессор, член-корреспондент, Национальной академии наук Беларуси, Республика Беларусь



**Тусупов Д.А.**, заведующий кафедрой информационных систем Евразийского национального университета им. Л.Н. Гумилева, доктор физико-математических наук, профессор, Республика Казахстан



**Кругликов С. В.**, генеральный директор Объединенного института проблем информатики Национальной академии наук Беларуси



**Шнейдеров Е.Н.**, проректор по учебной работе Белорусского государственного университета информатики и радиоэлектроники, кандидат технических наук, доцент, Республика Беларусь





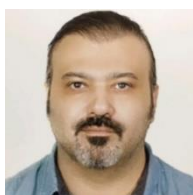
**Стемпичский В.Р.**, проректор по научной работе БГУИР, начальник научно-исследовательской части, научный руководитель НИЛ "Компьютерное проектирование микро- и наноэлектронных систем" кандидат технических наук, доцент, Республика Беларусь



**Шевченко Д.А.**, декан факультета управления Южного федерального университета (г. Ростов-на-Дону), кандидат экономических наук, доцент, Российская Федерация



**Тебуева Т.Б.**, заведующая кафедрой прикладной математики и компьютерной безопасности Северо-Кавказского Федерального университета, доктор физико-математических наук, доцент, Российская Федерация



**Косари А.**, старший консультант в области ИТ, кибер и информационной безопасности организации информационных технологий Ирана. Доцент кафедры компьютерной инженерии Международного кампуса Технологического университета Шарифа, кандидат технических наук, Иран

## ОРГАНИЗАТОРЫ КОНФЕРЕНЦИИ



Министерство образования Республики  
Беларусь



Учреждение образования  
«Белорусский государственный университет  
информатики и радиоэлектроники»



Национальная академия наук Беларуси



Объединенный институт проблем  
информатики



Ташкентский университет информационных  
технологий, Узбекистан



Южный федеральный университет, Российская  
Федерация



Северо-Кавказский федеральный университет,  
Российская Федерация



Евразийский национальный университет им. Л.Н.  
Гумилева



Технологический университет Шарифа, Иран

---

**СПОНСОР  
КОНФЕРЕНЦИИ**



**Облачный хостинг hoster.by**

Гибкий и надежный хостинг для интернет-проектов на любых платформах и CMS: от сайтов на WordPress до приложений на Java, Ruby, Python. Идеален для сайтов, которые переросли возможности обычного хостинга или требуют индивидуальных настроек.



## ОГЛАВЛЕНИЕ

<b>К.И. Котельников, В.Е. Буюков, А.Н. Марков</b> Применение big data в сфере экономики.....	11
<b>И.П. Кобяк</b> Производящая функция для вероятности пропуска ошибки при наблюдении векторов переходов.....	16
<b>Е.С. Дягель, Д.А. Фролова</b> Использование Big Data для формирования маркетинговых стратегий.....	24
<b>В.Ф.Алексеев, Г.А.Пискун, Д.В. Лихачевский, И.Н. Тонкович, Э.В. Асадчая, А.Д. Ларькин</b> Реализация регистрации участников на базе электронной площадки google forms с формированием онлайн базы данных.....	32
<b>С.А. Байчик, С.Н. Нестеренков, И.С. Тарасюк</b> Оптимизация обработки Big Data с помощью графических процессоров .....	44
<b>Е.И. Лешевич, П.В. Камлач, В.М. Бондарик, А.В. Чураков, Г.Д. Ситник, И.И. Ревинская</b> Математические вычисления с использованием графических процессоров по технологии CUDA .....	50
<b>Д.А. Фролова, Ю.А. Янукович</b> Использование веб-аналитики для решения бизнес-задач корпоративного сайта.....	59
<b>В.П. Домеников, А.Г. Сапёров, А.С. Строгова</b> Моделирование метода сжатия и обратного восстановления двоичного кода с использованием полиномиальных коррелированных функций .....	62
<b>П.А. Такунов, С.Н. Нестеренков, А.Н. Марков</b> Big Data в веб приложении по контролю выполненной работы в команде .....	67
<b>С.М. Боровиков, С.К. Дик</b> Использование программно-информационного комплекса Арион в it-образовательной среде.....	72
<b>Ван Там Лэ, С.М. Боровиков, С.С. Дик, А.В. Будник</b> Рекомендации по оценке и обеспечению надёжности планируемых к разработке прикладных компьютерных программ для информационных систем.....	78
<b>А.И. Лордкипанидзе, В.Д. Владимирцев, А.Н. Марков</b> Качество данных в Big Data .....	87
<b>С.С. Марковский, С.Н.Нестеренков</b> Использование больших данных в развитии современного земледелия .....	91
<b>А С.А. Павлюковец,А. Вельченко, У Синьсинь, А.А. Радкевич , Н.О. Савко</b> К вопросу управления мобильным роботом с колесами всенаправленного типа .....	94
<b>Н.В. Павлович, И.Н. Тонкович, В.Ф. Алексеев</b> Ключевые аспекты выбора СУБД для аналитики больших данных .....	103
<b>Ю.П. Плахотникова, Т.С. Ширяев, С.Н. Нестеренков</b> Data Analytics в чатах .....	109
<b>М.В. Реут, В.Д. Владимирцев, А.Н. Марков</b> Анализ социальных сетей с помощью Big Data: выявление трендов.....	113
<b>С.М. Сацук, С.В. Дробот, В.Н. Русакович</b> Информационная среда для оптимизации надзора в области обеспечения ядерной и радиационной безопасности при эксплуатации Белорусской АЭС.....	117
<b>С.А. Павлюковец, А.А. Вельченко, Ян Шисинь, Д.Ю. Чаплыгин, А.А. Радкевич, Н.О. Савко</b> Алгоритмика управления четырехколесными мобильными роботами .....	121

<b>Ю.П. Плахотникова, К.В. Спесивцев, С.Н. Нестеренков</b> Использование Big Data в гостиничном бизнесе .....	132
<b>Д.А. Сухарко, В.Д. Владимцев, А.Н. Марков</b> Преимущества использования Big Data в образовании и учебном процессе .....	135
<b>Д.В. Тавлуй, В.Ф. Алексеев, Г.А. Пискун</b> Веб-аналитика как инструмент для измерения ....	138
<b>В.С. Вашкевич, Л.Ф. Васковская, А.В. Гордиевич</b> Автоматизированная система идентификации объектов по радиолокационным и картографическим данным.....	145
<b>Н.А. Калабин, А.О. Чмутов, В.Д. Владимцев</b> Оценка доверия к различным информационным источникам и методы работы с ними в эпоху Big Data.....	154
<b>В.П. Корячко, А.В. Бакулев, М.А. Бакулева</b> Алгоритм нахождения оптимального пути в транспортной сети в условиях изменяемых параметров .....	161
<b>Д.Г. Готченя, О.Э. Осадчий, В.Д. Владимцев</b> Финансовый навигатор: концепт платформы для управления личными финансами на основе анализа больших данных.....	167
<b>Д.В. Тавлуй, В.Ф. Алексеев, Г.А. Пискун</b> Использование искусственного интеллекта в интернет-маркетинге .....	172
<b>Д.А. Брезин, С.Н. Нестеренков, А.Н. Марков</b> Использование технологий Big Data для анализа рынка недвижимости.....	178
<b>А.М. Тургунов</b> Применение Big Data в управление протезом сердца человека.....	181
<b>Е.Н. Шнейдеров, М.П. Батура, А. С. Терешкова, К.С. Крез</b> Система мониторинга показателей образовательного процесса для образовательных платформ на базе moodle LMS .....	188
<b>М. М. Гресик, С. Н. Нестеренков</b> Поиск визуально подобных изображений на основе машинного обучения .....	195
<b>Д.И. Альховик, В.Д. Владимцев, А.Н. Марков</b> Big Data: Проблемы и технологии .....	201
<b>Д.Д. Дронов, Д.Ю. Перцев</b> Интеллектуальный анализ данных для раннего обнаружения лесных пожаров .....	205
<b>У.П. Курилина, Д.А. Филон, М.С. Ильясова, Ф.В. Усенко, Л.Р. Коркин, А.М. Прудник</b> Игра дополненной реальности «Битва с роботами».....	210
<b>К.А. Кот</b> Составление рекомендаций для магазинов торговой сети на основе анализа потребительской корзины .....	215
<b>Т.А. Васяева, С.В. Хмелевой, А.С. Усова</b> Интеллектуальная система оптимизации потребления электроэнергии в жилом доме.....	220
<b>М.М. Татур, А.Д. Конилов, И.Н. Носырев</b> Принципы работы и сравнение характеристик точности спутниковых приемников.....	227
<b>В.А. Милентьев, А.М. Прудник, И.В. Андриаович, А.В. Воробей</b> Программная поддержка управления взаимодействием с клиентами в сфере аренды автомобилей .....	236
<b>М.В. Большакова, М.С. Ильясова, Ф.В. Усенко, Л.Р. Коркин, А.М. Прудник</b> Игра виртуальной реальности «Космонавт Бесконечности» .....	246
<b>Д.А. Кечик, И.Г. Давыдов</b> Плотности вероятностей межкомпонентных фазовых отношений.....	253
<b>К.С. Дик, И.И. Пилецкий</b> Алгоритмы поиска неисправностей солнечной электростанции на основе больших данных её мониторинга .....	259

<b>А.Г. Давыдовский, Л.Н. Воронцов, А.В. Пищова</b> Цифровой университет 4.0 как образовательная социотехническая экосистема для индустрии 4.0.....	269
<b>С.Н. Барсукевич, С.Н. Нестеренков, Д.В. Низовцов</b> BIG DATA в краудфандинге как средство прогнозирования успеха кампании .....	279
<b>Н.В. Павлов</b> Обучение специалистов-маркетологов работе с данными .....	282
<b>М.А. Гергенсон, С.Н. Нестеренков, И.С. Тарасюк</b> Использование технологии Big Data при построении системы управления мобильным роботом.....	287
<b>Д. И. Альховик, А.К. Горбачев, В.Д. Владимцев, С.А. Мигалевич</b> Развитие использования нейросетей для анализа эстетических предпочтений с применением Big Data	292
<b>И.С. Куис, М.И. Вашкевич</b> Использование различных семейств вейвлетов в задаче слияние медицинских изображений на основе дискретного вейвлет-преобразования .....	299
<b>А.О. Чаплинский, С.Н. Нестеренков, И.Г. Скиба</b> Использование технологии Big Data для отслеживания и контроля успеваемости учащихся.....	309
<b>Г.А. Трофимук, П.А. Архиреев, А.Н. Марков</b> Анализ Big Data в веб-приложениях для настольных игр .....	313
<b>Ж.Т. Усмонов, Т.Б. Джураев, А.Э. Кувнаков</b> Организация рациональной структуры управления перевозочным процессом в железнодорожном транспорте .....	318
<b>С.Н. Кардаш</b> Построение блочных разбиений систем булевых функций на основе задачи покрытия булевых матриц.....	326
<b>Г.А. Трофимук, П.А. Архиреев, А.Н. Марков</b> Анализ Big Data в Gamemobile .....	331
<b>Д.А. Кечик, И.Г. Давыдов</b> Статистический анализ межкомпонентных фазовых отношений.....	336
<b>С.Ю. Слюсарь, Д.С. Коваленко, М.С. Ильясова, Ф.В. Усенко, Л.Р. Коркин, А.М. Прудник</b> Работа алгоритма коллапса волновой функции на примере игры «tower of elevation».....	342
<b>А.Б. Никульшин, В.М. Бондарик, Н.В. Русина</b> Повышение производительности автоматизированной системы подачи заявлений и зачисления за счет масштабирования базы данных .....	348
<b>Д.В. Орлов, С.Н. Нестеренков, Д.А. Жалейко</b> Big Data как основа построения эффективной системы мониторинга пандемий на примере COVID-19.....	352
<b>В.Ю. Красовский, С.Н. Нестеренков, И.Г. Скиба</b> Разработка инструментальных средств для изучения и анализа уровня владения иностранным языком .....	357
<b>И.В. Лосик, С.Н. Нестеренков, Д.В. Низовцов</b> Стеганографические методы сокрытия .....	360
<b>L.I. Arkhipova</b> Composable business in digitaltransformation .....	364
<b>V.A. Vishnyakov, Yu ChuYue, Xia YiWei</b> Presentation and processing of data for the diagnostics of neurological diseases .....	369
<b>A.N. Osipov, O.Ch. Rolich, M.P. Batura, A.P. Kluev, V.D. Vladymtsev, S.A. Migalevich, E.R. Uglanova</b> Heart rate measurement algorithm in the system of remote monitoring of human condition .....	374
<b>R.M. Charapennikau</b> Deep learning based approach for vertebrae detection on spine X-RAY .....	379



УДК 004.021:004.75

## ПРИМЕНЕНИЕ BIG DATA В СФЕРЕ ЭКОНОМИКИ



**К.И. Котельников**  
Студент БГУИР  
kki199918@gmail.com



**В.Е. Буюков**  
Студент БГУИР  
vitya.buiukov@mail.ru



**А.Н. Марков**  
Старший преподаватель,  
магистр технических наук,  
заместитель начальника Центра  
информатизации и  
инновационных разработок БГУИР  
a.n.markov@bsuir.by

### **В.Е. Буюков**

Студент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники.

### **К.И. Котельников**

Студент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники.

### **А.Н. Марков**

Магистр технических наук, старший преподаватель кафедры ПОИТ, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Сегодня Big Data проникли во все сферы нашей жизни, в том числе и в экономику. С помощью анализа больших данных можно улучшить, дополнить старые алгоритмы и создавать новые экономические модели. Целью данной статьи является анализ роли больших данных в сфере экономики. Также для анализа больших данных, рассматривается фреймворк Hadoop.

**Ключевые слова:** Big Data, экономика, технологии, Hadoop.

### **Введение.**

В современном мире объемы данных растут с невиданной скоростью, и их анализ и использование стали ключевым фактором для развития различных отраслей экономики. Одной из наиболее актуальных технологий, позволяющих эффективно обрабатывать и анализировать огромные объемы данных, является Big Data. В данной статье рассмотрены возможности использования Big Data в экономике и их влиянии.

### **Актуальность.**

Применение Big Data в экономике позволяет более точно оценить риски и преимущества различных стратегий, оптимизировать производственные процессы, улучшать качество продукции и услуг, увеличивать эффективность маркетинговых кампаний, определять предпочтения и потребности потребителей, управлять ресурсами и бюджетами более эффективно, и многое другое.

Применение Big Data в экономике также актуально в свете стремительного развития цифровых технологий, таких как интернет вещей (IoT), искусственный интеллект (ИИ), облачные вычисления и другие, которые способствуют генерации и накоплению еще больших объемов данных. Эти технологии предоставляют новые возможности для сбора, хранения, анализа и использования данных с целью оптимизации бизнес-процессов и создания ценности.

### **Определение Big Data.**

Big Data – это огромные объемы данных, которые накапливаются в различных источниках, таких как социальные сети, интернет-сайты, мобильные устройства, датчики, транзакции, и другие, и характеризуются тремя основными аспектами – объемом, разнообразием и скоростью обработки. В экономике Big Data представляют собой ценный источник информации, который может использоваться для анализа рынков, потребительского поведения, финансовых операций, производственных процессов и других аспектов экономической деятельности [1].

### **История использования Big Data**

Одним из ранних примеров использования Big Data в экономике является анализ рынков и инвестиций. Финансовые институты и инвесторы используют огромные объемы данных, таких как цены акций, объемы торгов, экономические показатели и другие финансовые данные, для прогнозирования рынка, определения трендов и принятия решений о покупке и продаже акций, облигаций, валют и других финансовых инструментов.

Также Big Data находит применение в оптимизации производственных процессов, управлении цепями поставок, прогнозировании спроса, определении оптимальной ценовой политики, управлении рисками и многое другое. Большие объемы данных, собранных от производственных линий, логистических систем, складов, транспортных сетей и других источников, позволяют компаниям оптимизировать свою деятельность, снижать издержки, увеличивать эффективность и улучшать качество продукции и услуг. Компании используют данные о покупках, предпочтениях, интересах и демографических характеристиках потребителей, собранных из различных источников, для определения предпочтений потребителей, прогнозирования спроса на товары и услуги, улучшения маркетинговых стратегий и повышения уровня удовлетворенности клиентов [2].

### **Технологии Big Data в экономике.**

Наиболее распространенными технологиями Big Data, которые могут применяться в экономическом контексте, являются:

**Сбор и анализ больших объемов данных:** С использованием Big Data технологий, таких как Hadoop, Spark и других инструментов, организации могут собирать, хранить и анализировать огромные объемы данных из различных источников, таких как социальные сети, датчики, транзакции, клиентские данные и другие, для выявления тенденций, паттернов и корреляций в экономических процессах.

**Машинное обучение и аналитика данных:** Применение алгоритмов машинного обучения и аналитики данных на основе Big Data позволяет организациям выявлять скрытые закономерности, прогнозировать рыночные тенденции, оптимизировать бизнес-процессы, определять предпочтения и поведение клиентов, проводить анализ рисков и многое другое.

**Распределенные базы данных и облачные вычисления:** Технологии Big Data также позволяют хранить и обрабатывать данные в распределенных базах данных, а также использовать облачные вычисления для более эффективного управления большими объемами данных, анализа и обработки.

**Интернет вещей (IoT):** Большие объемы данных могут быть сгенерированы и собраны с помощью различных устройств IoT, таких как датчики, умные дома, автомобили и другие, и использоваться для оптимизации бизнес-процессов, производственной деятельности, логистики и других аспектов экономической деятельности.

**Анализ социальных медиа:** Социальные медиа являются богатым источником данных, и Big Data технологии могут использоваться для анализа мнений, настроений, трендов и поведения пользователей в социальных сетях, что может быть ценной информацией для маркетинговых исследований, продвижения продуктов и услуг.

**Анализ финансовых данных:** В экономике финансовые данные играют важную роль, и Big Data технологии могут быть использованы для анализа данных о финансовых рынках,

транзакциях, инвестициях, кредитной и банковской деятельности, что может помочь в принятии решений в области инвестиций, риск-менеджмента и финансового планирования.

**Прогнозирование и оптимизация:** Применение Big Data технологий позволяет организациям проводить прогнозирование и оптимизацию бизнес-процессов на основе данных, таких как спрос, предложение, цены, конкурентная активность и другие факторы, что может помочь в принятии более информированных и эффективных решений в экономической деятельности.

**Анализ данных клиентов:** Big Data технологии могут быть использованы для анализа данных о клиентах, их предпочтениях, поведении и покупательской активности, что может помочь в улучшении маркетинговых стратегий, персонализации предложений и повышении удовлетворенности клиентов.

**Анализ данных логистики и цепей поставок:** Big Data технологии могут быть применены для анализа данных о логистических процессах, цепях поставок и транспортной логистике, что может помочь в оптимизации логистических операций, снижении затрат и улучшении эффективности доставки товаров и услуг.

**Анализ данных о конкурентной среде:** Big Data технологии могут быть использованы для анализа данных о конкурентной среде, таких как данные о ценах, продуктах, маркетинговых активностях конкурентов, что может помочь в принятии решений по ценообразованию, маркетинговой стратегии и развитию конкурентоспособности организации [3].

В дальнейшем изучение технологии будет проходить на примере фреймворка Hadoop.

#### **Использование Hadoop.**

Множество компаний по всему миру используют Hadoop в своих бизнес-процессах. Некоторые из известных компаний, которые применяют Hadoop в своей деятельности, включают:

Amazon – одна из крупнейших компаний в области электронной коммерции, использует Hadoop для анализа данных о покупках, логистике, рекомендательных системах и многих других аспектах своего бизнеса.

Google – одна из ведущих компаний в области интернет-технологий, использует Hadoop для обработки и анализа данных, таких как поисковые запросы, рекламные платформы, аналитика веб-трафика и др.

IBM – одна из крупнейших IT-компаний в мире, использует Hadoop в своих решениях для анализа и обработки данных, таких как аналитика бизнес-процессов, аналитика клиентского опыта, аналитика рынка и др.

Walmart – одна из крупнейших розничных сетей в мире, использует Hadoop для анализа данных о продажах, логистике, предсказательной аналитике спроса, а также для оптимизации поставок и управления запасами.

Intel – одна из крупнейших компаний в области производства микропроцессоров, использует Hadoop для анализа данных, связанных с производством, тестированием и оптимизацией процессов производства.

#### **Структура Hadoop.**

Hadoop – это фреймворк для обработки больших данных, разработанный Apache Software Foundation. Он состоит из двух основных компонентов: Hadoop Distributed File System (HDFS) и Hadoop MapReduce.

**Hadoop Distributed File System (HDFS):** Это распределенная файловая система, предназначенная для хранения и обработки больших объемов данных на кластере серверов. HDFS разбивает файлы на блоки и распределяет их на различные узлы в кластере, обеспечивая отказоустойчивость и масштабируемость. HDFS разбивает файлы на блоки и распределяет их на различные узлы в кластере, обеспечивая отказоустойчивость и масштабируемость.

**Hadoop MapReduce:** Это модель программирования и выполнения, которая позволяет эффективно обрабатывать большие объемы данных на кластере Hadoop. MapReduce разделяет задачи на две основные операции: Map (отображение) и Reduce (сведение), которые выполняются параллельно на различных узлах кластера.



**Hadoop YARN (Yet Another Resource Negotiator):** Это система управления ресурсами, которая отвечает за управление ресурсами кластера, распределение задач на вычислительные узлы и мониторинг их выполнения. YARN позволяет более эффективно использовать ресурсы кластера и поддерживать множество различных рабочих нагрузок, таких как MapReduce, Apache Spark, Apache Hive и др.

**Hadoop EcoSystem:** Это набор дополнительных инструментов и технологий, разработанных на основе Hadoop, таких как Apache Hive, Apache Pig, Apache HBase, Apache Spark и другие. Эти компоненты расширяют функциональность Hadoop и позволяют более удобно и эффективно выполнять различные задачи обработки и анализа данных.

Рабочий процесс Hadoop обычно состоит из следующих шагов:

**Хранение данных в HDFS:** Данные разделяются на блоки и сохраняются в HDFS на различных узлах кластера. HDFS обеспечивает отказоустойчивое хранение данных, распределяя данные по узлам и создавая несколько реплик каждого блока данных.

**Обработка данных с помощью MapReduce:** Программа MapReduce разделяет обработку данных на два этапа - Map и Reduce. Эти этапы выполняются на распределенных узлах кластера.

– **Этап Map:** В этом этапе данные извлекаются из HDFS и обрабатываются на каждом узле отдельно. Результаты преобразования записываются в локальные файлы на каждом узле.

– **Этап Reduce:** В этом этапе результаты Map объединяются и обрабатываются для получения окончательных результатов. Reduce-функции выполняют агрегацию, фильтрацию, сортировку и другие операции над данными.

**Обработка результата:** Окончательные результаты обработки сохраняются в HDFS или передаются для дальнейшей обработки или анализа.

**Мониторинг и управление:** Hadoop предоставляет инструменты для мониторинга и управления кластером, такие как Hadoop YARN (Yet Another Resource Negotiator) и Hadoop MapReduce JobTracker, которые позволяют контролировать выполнение задач, оптимизировать использование ресурсов и обеспечивать отказоустойчивость [4].

### **Показатели эффективности**

**Twitter:** Благодаря использованию Hadoop для анализа и обработки миллиардов твитов ежедневно, Twitter смог повысить эффективность своих маркетинговых кампаний. Анализ метрик, таких как количество упоминаний, ретвитов, лайков и других показателей, позволил Twitter оптимизировать свои маркетинговые стратегии и повысить эффективность на 10%.

**eBay:** Использование Hadoop позволило eBay анализировать огромные объемы данных о продажах, клиентах и других метриках, что помогло оптимизировать его платформу. В результате, eBay смог улучшить процесс продаж и повысить свою эффективность на 15% [5].

### **Заключение.**

В заключение, использование Big Data в экономике имеет огромный потенциал для оптимизации бизнес-процессов, принятия обоснованных решений и повышения эффективности компаний. С помощью передовых технологий, таких как Hadoop, компании могут обрабатывать и анализировать огромные объемы данных, выявлять скрытые паттерны и тренды, а также прогнозировать будущие события.

### **Список литературы**

- [1] Gartner, Inc. "Big Data." Gartner IT Glossary [Электронный ресурс], Адрес: <https://www.gartner.com/en/information-technology/glossary/big-data>.
- [2] Mayer-Schönberger, V., & Cukier, K. (2013). Большие данные: революция, которая изменит то, как мы живем, работаем и думаем. Хоутон Миффлин Харкорт.
- [3] Доббс, Р., Роксбург, К., & Байерс, А. Х. (2011). Большие данные: следующий рубеж для инноваций
- [4] Документация Hadoop [Электронный ресурс], - Адрес: <https://hadoop.apache.org/>
- [5] Mayer-Schönberger, V., & Cukier, K. (2013). Big Data: A Revolution That Will Transform How We Live, Work, and Think. Houghton Mifflin Harcourt.

## **APPLICATION OF BIG DATA IN THE FIELD OF EDUCATION**

***K.I.Katselnikau***

*Student of the BSUIR*

***V.E. Buyukov***

*Student of the BSUIR*

***A.N.MARKOV***

*Senior lecturer of the department,  
Deputy head of the Center for  
Informatization and Innovative  
Developments*

*Department of Computer Science*

*Faculty of Computer Systems and Networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: kki199918@gmail.com*

**Abstract.** Today, Big Data has penetrated into all spheres of our life, including the economy. With the help of big data analysis, it is possible to improve, supplement old algorithms and create new economic models. The purpose of this article is to analyze the role of big data in the field of economics. Also for big data analysis, the Hadoop framework is being considered.

**Keywords:** Big Data, economics, technology, Hadoop.

УДК 519.624.2

## ПРОИЗВОДЯЩАЯ ФУНКЦИЯ ДЛЯ ВЕРОЯТНОСТИ ПРОПУСКА ОШИБКИ ПРИ НАБЛЮДЕНИИ ВЕКТОРОВ ПЕРЕХОДОВ



**И.П. Кобяк**  
доцент кафедры ЭВМ  
канд. техн. наук  
IPKobyak2012@mail.ru

### **И.П. Кобяк**

Работает в Белорусском государственном университете с 1982 г. Занимаемые должности инженер, ассистент, доцент кафедры ЭВМ. Защитил кандидатскую диссертацию в 1993 г. Область научных интересов: методы идентификации сообщений, проектирование спецпроцессоров.

**Аннотация.** В работе рассмотрен метод синтеза производящей функции для вероятности наблюдения векторов переходов в задачах анализа и синтеза контрольных кодов. Определено соотношение для общего случая данного параметра, соответствующее регистрации произвольного числа заданных пар событий. Полученное соотношение позволило выполнить сравнение уровней пропуска ошибки исследуемым методом с известными алгоритмами формирования контрольных кодов, такими как линейное сверточное кодирование и синтез оценок числа векторов состояний.

**Ключевые слова:** векторы переходов, субдинамические объекты, постобъекты, идентификация сообщений, вероятность пропуска ошибки, сигнатурный анализ.

### **Введение.**

В работах [1,2] выполнена постановка задачи исследования структурных компонентов  $r$ -разрядных последовательностей с точки зрения представления многомерных событий в виде сложных объектов длиной  $2j+i$ , где  $j$  - это число последовательных векторов переходов (ВП),  $i$  - длина постобъекта из векторов состояний (ВС) или частный случай функции правдоподобия. На основе вероятностного анализа в [1] было получено равенство для  $j_{mo}$  - математического ожидания числа последовательных ВП в выборке длиной  $n \rightarrow \infty$  в виде:

$$j_{mo} = \sum_{j=1}^{0,5n} j p^j \left( \sum_{j=1}^{0,5n} p^j \right)^{-1} = \frac{1}{1-p}. \quad (1)$$

Сравнительный анализ метода наблюдения ВП заданного вида (или лебеговской меры ВП) с сигнатурным анализом или методом наблюдения ВС показал, что в асимптотике принцип идентификации случайных процессов вероятностью ВП имеет преимущество перед известными методами синтеза точечных оценок [1].

Однако дополнительные исследования данной задачи позволили сделать вывод, что полученные соотношения, не учитывают ряда факторов, определяющих математический механизм связи векторов состояний в постобъектах, то есть в вариантах функции правдоподобия. Кроме того, в известных публикациях до настоящего времени нет точного описания вероятности ошибки наблюдения лебеговской меры ВП в интегральной форме. Такой формой является производящая функция (ПФ) для данного параметра, позволяющая на основе математических



преобразований в алгебре степенных рядов ответить на вопросы, сформулированные конкретными техническими приложениями.

Таким образом, вопрос получения точного описания интегральных параметров по результатам статистического синтеза оценок ВП (в конечной или бесконечной выборке) для задач практического использования является актуальным и рассматривается в представляемой работе.

#### Математическая постановка задачи.

При определении уровня вероятности ошибки в задаче наблюдения сложных вероятностных объектов будем считать, что  $k_\omega$  - это общее число ВП заданного вида, которые формируются в результате технического «дифференцирования» цифровой последовательности с «однополупериодным выпрямлением» с целью синтеза субдинамических объектов (или ВП).

При этом под динамикой процесса понимается синхронное формирование выходных векторов в некоторой аппаратной или информационной среде на рабочих частотах.

При формировании формулы для числа различных перестановок с повторениями сложных событий на  $n$  местах размещения, рассмотрим принцип возникновения дискретных объектов в составе искомой вероятности.

Так, каждый субдинамический объект  $k_{j,i} \in k_\omega$  с  $\mu$  единичными битами может быть образован  $3^{r-\mu}$  способами, что определяет общую вероятность наблюдения соответствующих событий равную:

$$p = \frac{3^{r-\mu}}{m^2}. \quad (2)$$

Однако, более детальный анализ вероятности (2), рассмотренной через призму соотношения (1), позволяет заключить, что данный параметр, с учетом преобразования  $j_{mo} = 1 + \xi$ , для задач с полиномиальным представлением функций, должен быть трансформирован в соотношение:

$$p \rightarrow \frac{1}{\xi + 1} p + \frac{\xi}{\xi + 1} \sum_{j=2}^{0,5n} p^j < \frac{1}{\xi + 1} p + \frac{\xi}{\xi + 1} p^2. \quad (3)$$

Данный подход позволяет учесть, не только общее число ВП в лебеговской мере, но и определить их количество для каждого члена вероятности (2) в полиномиальном разложении вероятностной функции.

В общем случае можно показать, что, если векторы переходов наблюдаются с вероятностью (2) при  $\mu \neq r$ , и максимальном  $p = \frac{3}{16}$  (в задачах наблюдения ВП), из равенства (1) имеем:

$$j_{mo} = \frac{m^2}{m^2 - 3^{r-\mu}} = \left[ 1 - \left( \frac{3}{4} \right)^r \frac{1}{3^\mu} \right]^{-1} < 2.$$

С увеличением  $r$  параметр  $m^2$  возрастает быстрее, чем значение  $3^{r-\mu}$ .

Таким образом, с увеличением разрядности процесса  $r$  математическое ожидание  $j_{mo}$  также стремится к единице, а пространство вероятности лебеговской меры ВП заданного вида может быть представлено практически только двумя элементами (3).

Общий случай представления сложных событий в виде суммы векторов  $2j + i$  позволяет

сформировать факториальные моменты для функции распределения вероятностей ошибки, используя принцип включения и исключения вида:

$$Q_{j,i} = \left[ (3^{r-\mu})^j m^i \sum_{s=0}^z (-1)^s C_{i-s}^s (3^{r-\mu})^s \frac{1}{m^{2s}} \right]^{k_{j,i}} C_{n-g}^{k_{j,i}} = \left[ (3^{r-\mu})^j m^i \sum_{s=0}^z C_{i-s}^s (-p)^s \right]^{k_{j,i}} C_{n-g}^{k_{j,i}}, \quad (4)$$

где  $n = f(2j + i)$  по всем разбиениям на части с числом  $j$  ВП и постобъектами  $i$ , параметр  $g = f(2j + i - 1)$  приводит длину выборки  $n$  к перестановкам сложных объектов, а параметр  $C_{n-g}^{k_{j,i}}$  - это биномиальный коэффициент, определяющий число перестановок векторов статистики  $k_{j,i}$  на  $n - g$  местах размещения, значение  $z = ]0, 5i[$  как показано в [2].

Таким образом, задача синтеза производящей функции с использованием моментов (4) состоит в записи общей формы произведения для всех  $j$  и  $i$  и переходе от произведения к сумме сложных объектов в степени  $n - g$ .

**Производящая функция для вероятности пропуска ошибки при регистрации векторов переходов.**

Одним из базовых моментов при синтезе производящей функции для вероятности пропуска ошибки  $P_{ifc} = Mo + P_{ifc}(gl)$  (где  $Mo = \frac{1}{2^{n+1}} \beta_{0,n}$  - мода) является скорректированная зависимость из [2], позволяющая записать произведение для гладкой части функции  $P_{ifc}(gl)$  в виде:

$$P_{ifc}(gl) = \sum_g \pi(g) \sum_{n-g} \frac{1}{m^n} (3^{r-\mu})^{\deg} \left( \sum_{j=1}^{0,5n-2n-2j-2} \sum_{i=1}^{0,5n} jk_{j,i} + \sum_{j=1}^{0,5n} jk_{j,n-2j} \right) m^{\deg} \left[ \sum_{j=1}^{0,5n-2n-2j-2} \sum_{i=1}^{0,5n} ik_{j,i} + \sum_{j=1}^{0,5n} (n-2j)k_{j,n-2j} \right] \times \\ \times \prod_{j=1}^{0,5n-2} \left[ \prod_{i=2}^{n-2j-2} \left( \frac{1}{2^{i+1}} \beta_{j,i} \right)^{k_{j,i}} \left( \frac{1}{2^{i+1}} \beta_{j,n-2j} \right)^{k_{j,n-2j}} \right] \left[ \frac{1}{2^3} \beta_{\frac{n-2}{2},2} \right]^{\frac{k_{n-2}}{2}} \frac{(n-g)!}{\prod_{j=1}^{0,5n-2n-2j-2} \prod_{i=1}^{0,5n} k_{j,i}! \prod_{j=1}^{0,5n} k_{j,n-2j}!}, \quad (5)$$

где  $ifc$  - interaction function correlation,  $\pi(g)$  - композиция разбиений числа  $g$  в разбиениях числа

$$n = \sum_{j=1}^{0,5n-2n-2j-3} \sum_{i=1}^{0,5n} (2j + i) k_{j,i} + \sum_{j=1}^{0,5n} n k_{j,n-2j}$$

на целые части при  $k_{j,n-2j-1} = 0, \forall j \neq 0, 5n$ .

Составляющая  $n - g$  факториального момента в (5) с учетом ортонормирующего свойства  $g$  может быть приведена к равенству:

$$n - g = n - \sum_{j=1}^{0,5n-2n-2j-3} \sum_{i=1}^{0,5n} (2j + i - 1) k_{j,i} - \sum_{j=1}^{0,5n} (n - 1) k_{j,n-2j} = \sum_{j=1}^{0,5n-2n-2j-3} \sum_{i=1}^{0,5n} k_{j,i} - \sum_{j=1}^{0,5n} k_{j,n-2j}. \quad (6)$$

Функция правдоподобия  $\frac{1}{2^{i+1}} \beta_{j,i}$  в формуле (5) следует из (4) в соответствии

с результатом [3]:

$$\sum_{s=0}^z C_{i-s}^s (-p)^s = \frac{1}{2^{i+1}} \beta_{j,i}.$$

Далее, при переходе от произведения (5) к сумме элементов производящей функции необходимо помнить, что в скобочной форме ПФ будут отсутствовать в явном виде значения  $k_{j,i}$ , которые в преобразовании (6) переходят в степень ПФ.

Кроме того, необходимо помнить, что в условиях эксперимента некоторая частная статистика  $k_{j,i}$  может быть получена равной нулю. Данный факт приводит к изменению показателя степени  $n - g$ . Однако, в сумме моментов в ПФ принцип исключения слагаемых

$\frac{1}{2^{i+1}} \beta_{j,i}$  при  $k_{j,i} = 0$  в явном виде отсутствует. Таким образом, для всех слагаемых в ПФ необходимо вводить дополнительный коэффициент  $\sigma_{j,i} \in \left\{0, \frac{n}{2}\right\}$ , который будет указывать на факт

отсутствия или наличия каждого конкретного слагаемого в функции, а также на число повторений соответствующего факториального момента в выборке.

Следовательно, с учетом (1) можем записать соотношение для ПФ с параметрами  $j = 1$  и  $2$  в виде:

$$P_{jfc} < Mo + \sum_g \pi(g) \sum_{j=1}^2 \left[ A \left( \sigma_{j,i} p^j e^{jt} x_j^1 + \sum_{i=2}^{n-2j-2} \sigma_{j,i} p^j e^{jt} x_j^i \frac{1}{2^{i+1}} \beta_{j,i} + \right. \right. \\ \left. \left. + \sigma_{j,n-2j} p^j e^{jt} x_j^{n-2j} \frac{1}{2^{n-2j+1}} \beta_{j,n-2j} \right) \right]^{n-g}, \quad (7)$$

где коэффициент  $A$  в соответствии с (3) равен:

$$A = \left[ (2-j) + \xi(j-1) \right] \frac{1}{1+\xi}.$$

Запишем теперь общую форму ПФ, преобразуя произведение (5) в сумму, аналогично методике приведенной в [4] и [5], тогда:

$$\begin{aligned}
 P_{ifc} < \frac{1}{2^{n+1}} \beta_n + \sum_g \pi(g) \left\{ \frac{1}{\xi+1} \left( \sigma_{1,1} x_1^1 p e^t + \sum_{i=2}^{n-4} \sigma_{1,i} x_1^i p e^t \frac{1}{2^{i+1}} \beta_{1,i} + \sigma_{1,n-2} x_1^{n-2} p e^t \frac{1}{2^{n-1}} \beta_{1,n-2} \right) + \right. \\
 + \frac{\xi}{\xi+1} \left[ \left( \sigma_{2,1} x_2^1 p^2 e^{2t} + \sum_{i=2}^{n-6} \sigma_{2,i} x_2^i p^2 e^{2t} \frac{1}{2^{i+1}} \beta_{2,i} + \sigma_{2,n-4} x_2^{n-4} p^2 e^{2t} \frac{1}{2^{n-3}} \beta_{2,n-4} \right) + \right. \\
 + \left( \sigma_{3,1} x_3^1 p^3 e^{3t} + \sum_{i=2}^{n-8} \sigma_{3,i} x_3^i p^3 e^{3t} \frac{1}{2^{i+1}} \beta_{3,i} + \sigma_{3,n-6} x_3^{n-6} p^3 e^{3t} \frac{1}{2^{n-5}} \beta_{3,n-6} \right) + \dots + \\
 + \left( \sigma_{\frac{n-6}{2},1} x_{\frac{n-6}{2}}^1 p^{\frac{n-6}{2}} e^{\frac{n-6}{2}t} + \sum_{i=2}^4 \sigma_{\frac{n-6}{2},i} x_{\frac{n-6}{2}}^i p^{\frac{n-6}{2}} e^{\frac{n-6}{2}t} \frac{1}{2^{i+1}} \beta_{\frac{n-6}{2},i} + \sigma_{\frac{n-6}{2},6} x_{\frac{n-6}{2}}^6 p^{\frac{n-6}{2}} e^{\frac{n-6}{2}t} \frac{1}{2^7} \beta_{\frac{n-6}{2},6} \right) + \\
 + \left( \sigma_{\frac{n-4}{2},1} x_{\frac{n-4}{2}}^1 p^{\frac{n-4}{2}} e^{\frac{n-4}{2}t} + \sum_{i=2}^2 \sigma_{\frac{n-4}{2},i} x_{\frac{n-4}{2}}^i p^{\frac{n-4}{2}} e^{\frac{n-4}{2}t} \frac{1}{2^{i+1}} \beta_{\frac{n-4}{2},i} + \sigma_{\frac{n-4}{2},4} x_{\frac{n-4}{2}}^4 p^{\frac{n-4}{2}} e^{\frac{n-4}{2}t} \frac{1}{2^5} \beta_{\frac{n-4}{2},4} \right) + \\
 \left. \left. + \sigma_{\frac{n-2}{2},2} x_{\frac{n-2}{2}}^2 p^{\frac{n-2}{2}} e^{\frac{n-2}{2}t} \frac{1}{2^3} \beta_{\frac{n-2}{2},2} + \sigma_{\frac{n}{2},0} x_{\frac{n}{2}}^0 p^{\frac{n}{2}} e^{\frac{n}{2}t} \right] \right\}^{n-g}. \tag{8}
 \end{aligned}$$

В сокращенной форме данное равенство принимает вид:

$$\begin{aligned}
 P_{ifc} < \frac{1}{2^{n+1}} \beta_n + \sum_g \pi(g) \left[ \sum_{j=1}^{0,5n-2} \frac{A}{1+\xi} \left( \sigma_{j,1} x_j^1 p^j e^{jt} + \sum_{i=2}^{n-2j-2} \sigma_{j,i} x_j^i p^j e^{jt} \frac{1}{2^{i+1}} \beta_{j,i} \right) + \right. \\
 \left. + \frac{A}{1+\xi} \sigma_{\frac{n-2}{2},2} x_{\frac{n-2}{2}}^2 p^{\frac{n-2}{2}} e^{\frac{n-2}{2}t} \frac{1}{2^3} \beta_{\frac{n-2}{2},2} + \frac{A}{1+\xi} \sigma_{\frac{n}{2},0} x_{\frac{n}{2}}^0 p^{\frac{n}{2}} e^{\frac{n}{2}t} \right]^{n-g}, \tag{9}
 \end{aligned}$$

где  $A=1$ , при  $j=1$  и  $A=\xi$  при  $j>1$ .

Полученные соотношения (8) и (9) позволяют достаточно точно вычислить вероятность пропуска ошибки для конкретных реализаций многомерных последовательностей. Однако, в задаче расчета площади под интегральной кривой функции распределения вероятностей пропуска ошибки в соотношениях для ПФ коэффициенты  $\sigma_{j,i}$  могут быть взяты равными единицы, что позволяет суммировать все комбинаторные моменты, характеризующие частные вероятности ошибки, но на ограниченных множествах  $j$  и  $i$ . При этом для частного  $j=1$  на основании равенства (9) можем записать:

$$P_{ifc} \approx \frac{1}{2^{n+1}} \beta_n + \left( x_1^1 p e^t + \sum_{i=2}^{i_{m1}} x_1^i p e^t \frac{1}{2^{i+1}} \beta_{1,i} \right)^{n-g}, \tag{10}$$

где  $i_{m1}$  - это максимальное значение числа комбинаторных моментов в функции правдоподобия, удовлетворяющее равенству:

$$i_{m1} = -\frac{5}{2} + \frac{1}{2} \sqrt{8n+25}.$$

Данное соотношение может быть получено путем анализа закона суммирования длин

сложных объектов  $2+i$  с использованием методики подсчета элементов арифметической прогрессии. При этом:

$$n \approx \frac{(2+i_{m1})(3+i_{m1})}{2} - (1+2).$$

откуда и следует значение  $i_{m1}$ .

В равенстве (10) показатель степени

$$n-g = \sum_{i=1}^{i_{m1}} (2+i)k_{1,i} - \sum_{i=1}^{i_{m1}} (2+i-1)k_{1,i} = \sum_{i=1}^{i_{m1}} k_{1,i}, \quad k_{1,i} = 1, \quad n-g = i_{m1}.$$

При  $j=2$ , что практически полностью удовлетворяет критерию (1), из (7) имеем соотношение:

$$P_{ifc} < \frac{1}{2^{n+1}} \beta_n + \sum_{j=1}^2 \left[ \frac{A}{1+\xi} \left( p^j e^{jt} x_j^1 + \sum_{i=2}^{i_{mj}} p^j e^{jt} x_j^i \frac{1}{2^{i+1}} \beta_{j,i} \right) \right]^{n-g}, \quad (11)$$

где  $A=1$ , при  $j=1$  и  $A=\xi$  при  $j>1$ .

В данном случае параметр:

$$i_{m2} = -\frac{9}{2} + \frac{1}{2} \sqrt{8n+81}.$$

Соответственно степень:

$$n-g = \sum_{i=1}^{i_{m1}} k_{1,i} + \sum_{i=1}^{i_{m2}} k_{2,i}, \quad k_{1,i} = k_{2,i} = 1, \quad n-g = i_{m1} + i_{m2}.$$

Аналогично соотношениям (10) и (11), могут быть получены равенства для всех  $j$ , что, однако, не дает практически ничего нового для вероятности (11). Хотя уменьшение уровня указанного параметра с учетом приведения множителя  $A$  к полиномиальной форме с использованием (1) и (3) очевидно.

**Асимптотика и сравнительный анализ метода наблюдения ВП с другими классическими алгоритмами свертки (пример).**

В случае расчета асимптотики с использованием вероятности (2) можно записать равенство для ПФ (5) в упрощенном виде:

$$P_{ifc}(\infty) = \left( p^{1+\xi} \frac{1}{2^{i_{\infty}+1}} \beta_{1,i_{\infty}} \right)^{n-g_{\infty}}. \quad (12)$$

При максимальной вероятности наблюдения ВП равной  $\frac{3}{16}$  (для ВП любого вида) из соотношения (1) получаем:

$$p^{j_{mo}} = \left( \frac{3}{16} \right) \deg \left( \frac{16}{13} \right) = 0,1274.$$



Асимптотическая длина векторов функции правдоподобия для выбранной вероятности будет равна

$$i_{\infty} = \frac{1-2p^{j_{mo}}}{p^{j_{mo}}} = 5,848.$$

Тогда:

$$g_{\infty} = np^{j_{mo}}(i_{\infty} + 1) = 0,8726n.$$

$$n - g_{\infty} = 0,1274n.$$

Параметр  $\frac{1}{2^{i_{\infty}+1}} \beta_{i_{\infty}}$  рассчитывается по общей формуле:

$$\frac{1}{2^{6,848}} \beta_{1,5,848} = \frac{1}{2^{6,848}} \frac{1}{\sqrt{1-4 \cdot 0,1274}} \left[ \left(1 + \sqrt{1-4 \cdot 0,1274}\right)^{6,848} - \left(1 - \sqrt{1-4 \cdot 0,1274}\right)^{6,848} \right] = 0,4697.$$

Соответственно из (12) при этом имеем  $P_{ifc}(\infty) = (0,69855)^n$ .

Аналогичный показатель для вероятности ошибки при сигнатурном анализе в асимптотике имеет вид:

$$P_{msa} = \frac{1}{rn}.$$

Составляя теперь отношение вероятностей при  $r = 2$ :

$$\frac{P_{ifc}(\infty)}{P_{msa}} = (0,69855)^n 2n$$

получаем  $P_{ifc}(\infty) < P_{msa}$  практически во всем диапазоне значений  $n$ .

Заметим, однако, что с уменьшением вероятности  $p$  преимущество метода наблюдения лебеговской меры ВП уменьшается за счет факта  $P_{ifc}(\infty) \rightarrow (1-\Delta)^n$ , где  $\Delta$  - весьма малая величина. Тем не менее, в асимптотике неравенство  $P_{ifc}(\infty) < P_{msa}$  сохраняется всегда.

При сравнении полученного результата (12) с вероятностью ошибки, порождаемой методом наблюдения ВС можем записать равенство:

$$\frac{P_{ifc}(\infty)}{P_{cvc}} = (0,69855)^n \sqrt{\frac{3}{8} \pi n}$$

где  $cvc$  – *condition vector counting*, что также говорит о факте  $P_{ifc}(\infty) < P_{cvc}$ .

Таким образом, для асимптотики характерны неравенства  $P_{ifc}(\infty) < P_{msa} \leq P_{cvc}$ .

### Заключение.

В представленной работе получены следующие результаты:

1) на основе вероятностного анализа и равенства для  $j_{mo}$  - математического ожидания последовательных пар сложных событий в выборке (1) определено представление вероятности

наблюдения лебеговской меры ВП в виде двучлена (3) из полиномиального представления параметра;

2) получено произведение для гладкой части функции (5), определяющее вероятностную составляющую для числа перестановок векторов статистики  $k_{j,i}$  на  $n - g$  местах размещения с повторениями;

3) выполнен переход от произведения, характеризующего вероятность ошибки  $P_{ifc}$  к полиному производящей функции (8)-(9);

4) для вычисления площади под кривой вероятности пропуска ошибки получены соотношения (10) и (11), определяющие суммирование всех комбинаторных моментов, характеризующих частные вероятности ошибки на ограниченных множествах  $j$  и  $i$ ;

5) сравнительный анализ метода наблюдения лебеговской меры ВП с алгоритмами линейной свертки и СВС показал, что в асимптотике принцип идентификации случайных процессов вероятностью наблюдения ВП имеет явное преимущество перед известными алгоритмами синтеза точечных оценок.

### Список литературы

[1] Кобяк И.П. Асимптотика для вероятности пропуска ошибки при наблюдении векторов переходов. В кн.: BIG DATA и анализ высокого уровня 2021, BIG DATA and advanced analytics 2021: сборник материалов 7-й международной научно-практической конференции, Минск, 19-20 мая 2021 г. / Белорусский государственный университет информатики и радиоэлектроники, Минск, 2021. С. 328-335.

[2] Кобяк И.П. О точном равенстве для вероятности пропуска ошибки при наблюдении векторов переходов. В кн.: BIG DATA и анализ высокого уровня 2022, BIG DATA and advanced analytics 2022: сборник материалов 8-й международной научно-практической конференции, Минск, 11-12 мая 2022 г. / Белорусский государственный университет информатики и радиоэлектроники, Минск, 2022. С. 312-319.

[3] Риордан Дж. Комбинаторные тождества. Пер. с англ. – М.: Наука. Глав. ред. физ.-мат. литературы, 1982. 255 с.

[4] Кобяк И.П. Производящая функция для распределения статистик автокорреляционной функции // Электрон. моделирование. 2010. Т 32. – №2. – С. 61–76.

[5] Кобяк И.П. Производящая функция для вероятности пропуска ошибки при наблюдении двух векторов переходов. В кн.: Информационные технологии и системы 2022 (ИТС 2022), Information Technologies and Systems 2022 (ITS 2022): материалы международной научной конференции, Минск, 23 ноября 2022 г. / Белорусский государственный университет информатики и радиоэлектроники, Минск, 2022. С. 35-36.

## GENERATING FUNCTION FOR ERROR MISS PROBABILITY WHEN OBSERVING TRANSITION VECTORS

*I.P. Kobiak*

*PhD, Associate Professor, Chair of ECM, BSUIR.*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*

*E-mail: IPKobyak2012@mail.ru*

**Abstract.** The work considers the method of synthesis of the generating function for probability of occurrence of transition vectors in problems of analysis and synthesis of control codes. The ratio for the total case of this parameter is determined, corresponding to the registration of an arbitrary number of specified pairs of events. The obtained ratio made it possible to compare the error missing levels by the test method with known control code generation algorithms, such as convolution linear coding and observation of state vectors.

**Keywords:** transition vectors, sub dynamic objects, post-objects, identification of co-communications, probability of missing an error, signature analysis.

УДК 339.138

## ИСПОЛЬЗОВАНИЕ BIG DATA ДЛЯ ФОРМИРОВАНИЯ МАРКЕТИНГОВЫХ СТРАТЕГИЙ



**Е.С. Дягель**

Студентка 3 курса специальности «Электронный маркетинг» инженерно-экономического факультета БГУИР  
lizadziahel@gmail.com



**Д.А. Фролова**

Преподаватель кафедры экономики БГУИР  
frolova.profstud@gmail.com

### **Е.С. Дягель**

Окончила Гимназию N.2 г. Бреста. Область научных интересов связана с особенностями управления бизнес-процессами в мультиязычной цифровой среде, организацией процессов контент-маркетинга, построением стратегии интеллектуального лидерства (thought leadership strategy) при осуществлении маркетинговых коммуникаций B2B брендами в сети Интернет.

### **Д.А. Фролова**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с изучением механизмов влияния маркетинговых возможностей на формирование и развитие конкурентного бизнеса, нейромаркетинг.

**Аннотация.** Появление больших данных произвело революцию в понимании и взаимодействии компаний со своими клиентами. Посредством анализа огромных объемов информации, собранной из различных источников, большие данные позволяют получить ценные сведения о моделях поведения, предпочтениях и потребностях потребителей. Однако важно помнить о вопросах безопасности больших данных, т.к. это играет важную роль при формировании доверительных отношениях между клиентами и компанией.

**Ключевые слова:** большие данные, Big Data, data-driven маркетинг, анализ данных, оптимизация бизнес-процессов, маркетинговая аналитика.

### **Введение.**

Термин *Big Data* (большие данные) представляет собой большие объемы накопленных данных и технологии вычисления, хранения и интерпретации этих данных. Под большими данными подразумеваются технологии обработки неструктурированных и структурированных данных большого объема для получения понятных обычному человеку итогов [1].

В маркетинге *Big Data* используют для поддержки принятия управленческих решений, анализа поведения потребителей и определения тенденций в их покупательском поведении.

Клиент - ключевой «игрок» во всем маркетинге. Поэтому понимание того, для кого компания создает свои продукты или услуги, влияет на то будет ли успешна, выбранная маркетинговая стратегия.

Результаты, полученные на основе сбора и анализа больших данных, способны помочь специалистам в области маркетинга и коммуникаций создать наиболее точный портрет целевой аудитории, выяснить соответствует ли продукт или услуга тем требованиям и решениям, которые ожидает клиент; какие ценности у нашей целевой аудитории; помогает предсказать реакцию потребителя на новые предложения.

### **Big Data и продвижение.**

Изучив исследования и существующие отчеты о рынке, проанализировав статистику поисковых запросов и интересов пользователей интернета, компания будет предлагать продукт именно там, где его купят, может оптимизировать стратегии взаимодействия с лидами и клиентами, создавать такие ключевые сообщения, на которые будет реагировать аудитория.

Благодаря анализу существующих больших данных у компании формируется понимание того, что именно люди ищут, как они это делают, что они спрашивают, какие темы их сейчас

интересуют.

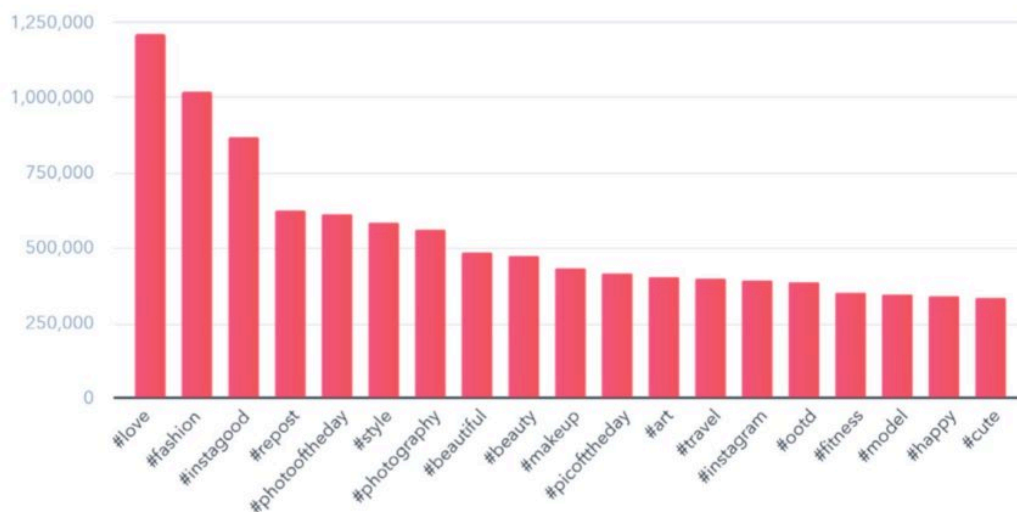


Рисунок 1. Рейтинг самых популярных хэштегов в Instagram

Так на примере графика вовлеченности по хэштегам в Instagram, представленном на рисунке 1, мы можем предположить, что если в своей публикации в данной социальной сети компания будет использовать хэштег #love, то охват ее публикации будет выше, за счет перехода на нее по популярному хэштегу.

Использование различных сервисов для аналитики и определения популярности поисковых запросов, трендов в интернете (*Google Analytics, Yandex Metrica, Convead, Serpstat, Яндекс Wordstat, Google Trends* и др.) позволяет получить полезную информацию о конкурентах компании и о пользователях сети, их запросах, присутствии в различных социальных сетях – это возможность оценить емкость рынка, понять необходимо ли компании присутствовать в поиске, стоит ли настраивать SEO продвижение, в каком формате, по какому направлению стоит двигаться, при создании контента.

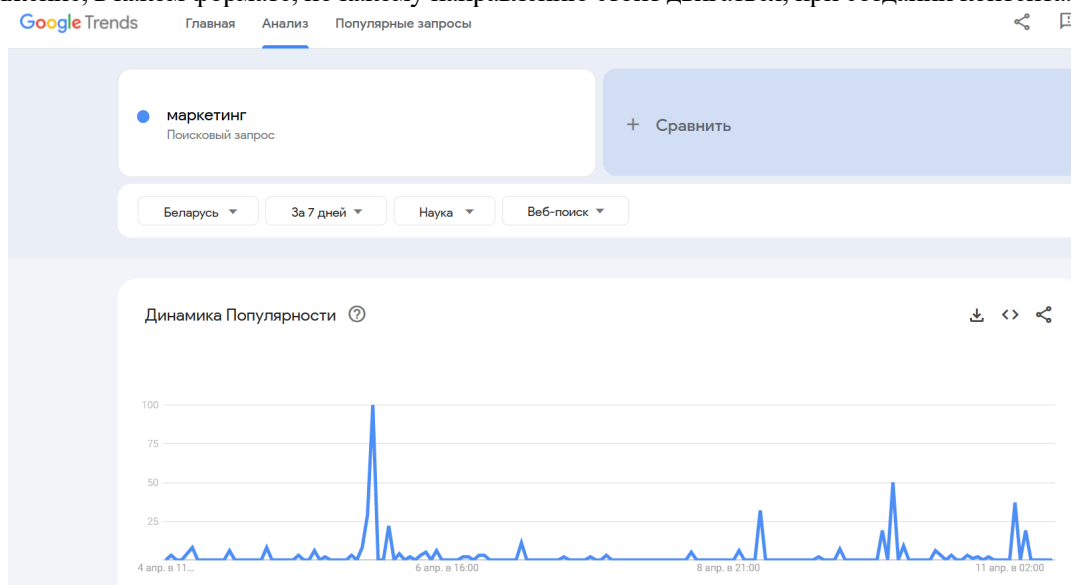


Рисунок 2. Пример анализа данных по запросу «маркетинг» в Google Trends

### Понятие data-driven маркетинга.

Маркетинг, основанный на данных, в английском языке получил название *data-driven marketing*. Принцип такого маркетинга заключается в том, что компании строят стратегии продвижения своих товаров и услуг, основываясь на выводах, полученных из анализа Big Data.

На сегодняшний день источников данных в маркетинге великое множество. К ним относятся данные клиентов (*customer data*):

- демографическая информация, включающая пол, возраст, местоположение и т.д.;
- данные об операциях (*transactional data*): информация о покупках и транзакциях клиентов, включающая купленные покупки, частоту транзакций и сумму потраченных средств;
- данные веб-аналитики, полученные с помощью инструментов отслеживания веб-сайтов Google Analytics, Яндекс.Метрика и прочих: информация о времени, проведенном на сайте, просмотренных страницах и их кликабельности;
- данные социальных медиа: вовлеченность пользователей, анализ настроений (*sentiment analysis*) и информация социального прослушивания (*social listening*);
- данные третьих лиц, или сторонние данные (*third-party data*): информация, полученная от сторонних поставщиков, таких как компании по исследованию рынка или брокеры данных;
- данные так называемой маркетинговой автоматизации (*marketing automation data*): информация, сгенерированная программными средствами маркетинговой автоматизации, такими как *HubSpot*, *InfusionSoft*, *SemRush* и прочими.

### Применение Big Data в сфере управления взаимоотношения с клиентами.

Специалисты в области маркетинга от вопроса «Где взять данные?» переходят к вопросу «Как интерпретировать большой объем данных и использовать эти данные для увеличения прибыли?». Согласно исследованию 2021 года *IBM Institute for Business Value*, 53% сферы применения Big Data приходится на сферу клиентского сервиса (рисунок 3) [3].

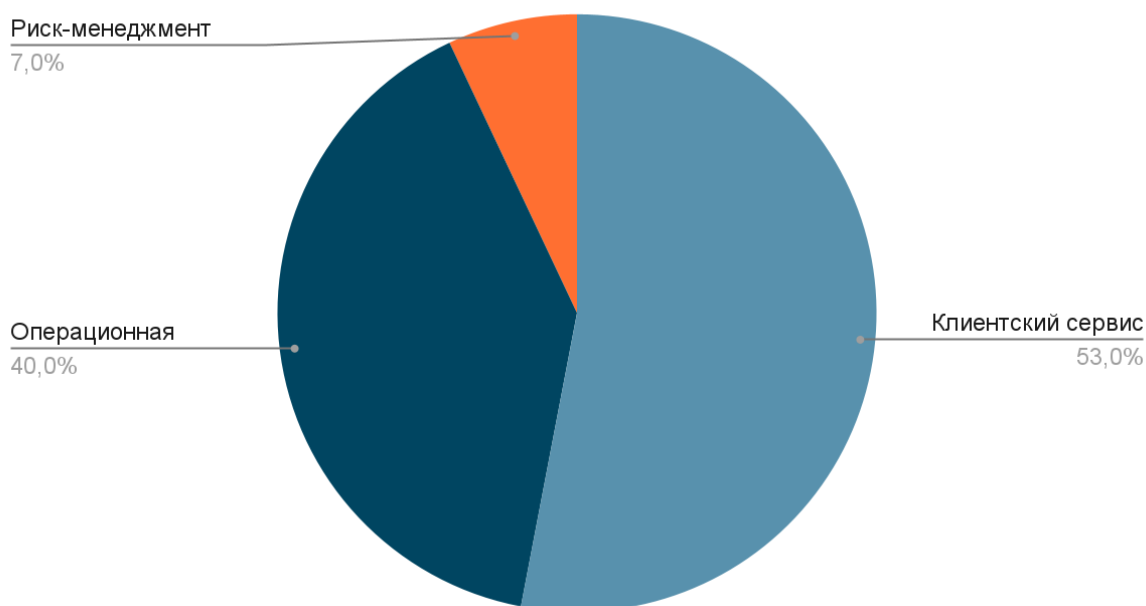


Рисунок 3. Сферы применения больших данных

Сегодня в условиях жесткой конкуренции на рынке, клиентоориентированный подход является важной составляющей при формировании маркетинговой стратегии. Исходя из диаграммы можно предполагать, что взаимоотношения с клиентами и дальше будут играть важнейшую роль в экономической и маркетинговой стратегиях компаний.

Поэтому для специалистов, отвечающих за *data-driven* маркетинг, важно грамотно использовать большой объем данных для наиболее полного удовлетворения потребностей клиентов.

Технологии получения данных (*data mining*) и их анализ дают возможность лучше понимать целевую аудиторию, выявлять неочевидные, но полезные на практике закономерности потребительского поведения, которые скрыты в больших объемах данных, а алгоритмы машинного обучения (*machine learning*) используются маркетологами и специалистами в области аналитики данных для прогнозирования поведения потребителей с использованием статистических методов и создания персонализированных предложений на основе построенного прогноза.

Маркетологи и дата-аналитики, работающие в сфере управления взаимоотношений с клиентами, используют ряд инструментов для работы с *Big Data*: *NoSQL Databases*, *BigQuery*, *MapReduce*, *Hadoop*, *WibiData*, *Skytree* – все они предназначены для сбора большого объема данных в одном месте, а такие платформы, как *Good Data*, *Power BI* от *Microsoft*, *Tableau Software*, *Toucan Toco*, *Zoho Analytics* помогают визуализировать собранные данные.

В итоге наглядно представленная на платформах информация, которая ранее была несистематически распределена в различных источниках (инструменты веб-аналитики, данные социальных медиа, и прочее) дает маркетологам возможность обнаружить паттерны поведения потребителей: например, с помощью упомянутых инструментов маркетолог может в режиме реального времени следить за закономерностями, которые прослеживаются в истории покупок, истории поиска, истории «лайков» как индивидуального потребителя, так и сегмента целевой аудитории.

Подробнее алгоритм работы с данными и целевые действия, которые могут быть предприняты для удовлетворения потребностей клиентов в цифровой среде на основе выявленных с помощью *Big Data* моделей их поведения, представлены на рисунке 4.

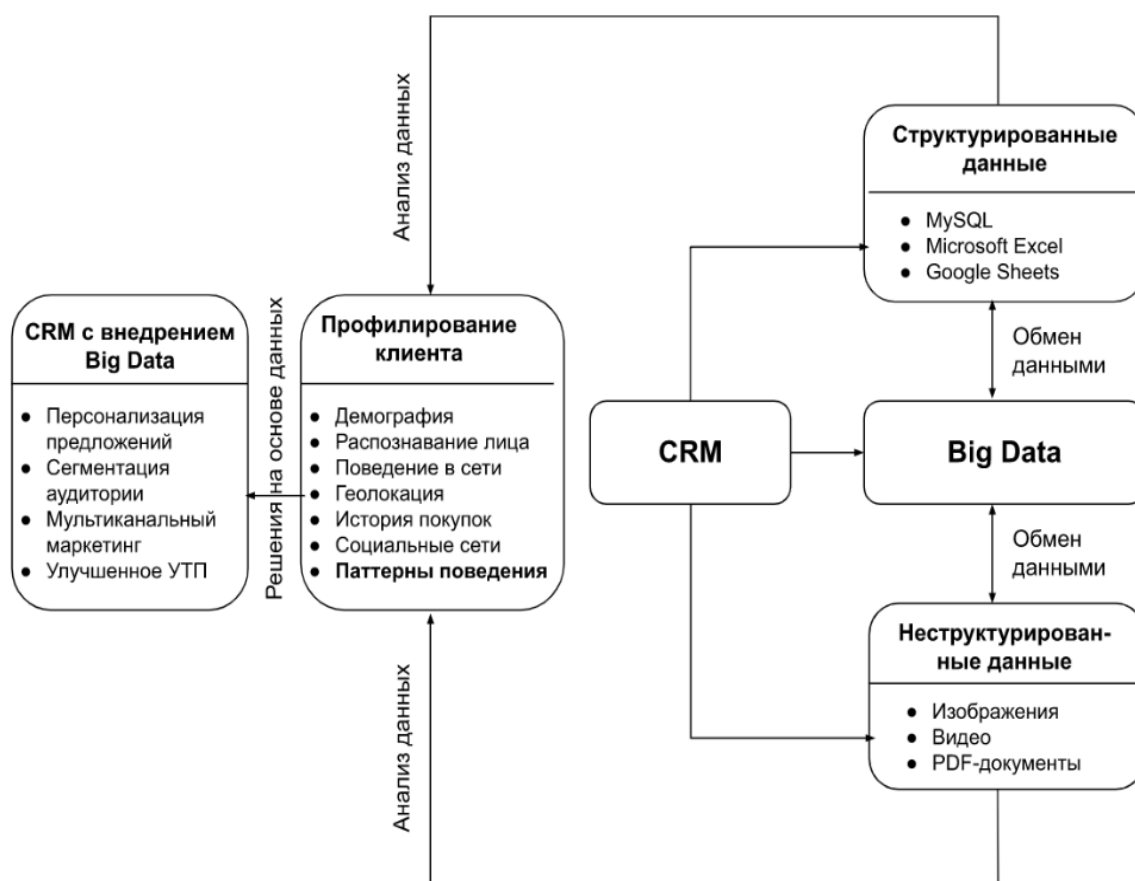


Рисунок 4. Внедрение *Big Data* в систему взаимоотношений с клиентами *CRM* [4]

Эффективное использование *Big Data* для персонализации клиентского опыта объясняет успех ряда зарубежных и отечественных компаний. Так, например, платформа электронной коммерции *Amazon*



использует технологии анализа данных для того, чтобы персонализировать рекомендации для своих пользователей в соответствии с их историей поиска и покупок.

Алгоритм стриминговой платформы *Netflix* собирает огромные объемы данных для того, чтобы анализировать личные предпочтения зрителя и предлагать ему фильмы, сериалы и телевизионные шоу с указанием процентного совпадения с его вкусами (например, «Совпадение: 86%»).

Российская технологическая компания Яндекс, предлагающая поисковую систему, электронную почту и онлайн-рекламу, использует анализ *Big Data* для предоставления персонализированных результатов поиска, контекстной рекламы, и индивидуальных рекомендаций контента.

Технологии *Big Data* используются и белорусскими ИТ-компаниями: так, белорусская компания по разработке программного обеспечения и ИТ-консалтингу *Epam Systems* использует анализ данных для создания специализированных веб-сайтов, мобильных приложений и прочих цифровых продуктов, которые учитывают потребности и предпочтения каждого отдельного пользователя.

Несмотря на широкое использование крупными корпорациями *Big Data* для оптимизации алгоритмов рекомендаций и персонализации своих предложений, остаются не до конца исследованными механизмами, которые помогут эффективно связать *Big Data* с маркетингом в предприятиях среднего и малого бизнеса [5]. Немногочисленные эмпирические исследования показывают, что команды маркетинга в небольших предприятиях (50-249, 250-499 сотрудников) еще не сполна используют потенциал больших данных для оптимизации бизнес-процессов и, в частности, отношений с клиентами.

Так, в 2021 агентством, специализирующимся на маркетинговых исследованиях в Китае, был проведен опрос 11 532 компаний малого и среднего бизнеса из следующих секторов: строительство, одежда, оборудование для дома, автомобили и аксессуары к автомобилям, электроника (рисунок 3). По результатам опроса была выявлена корреляция  $R^2 = 0,22$  между внедрением процесса маркетинговой аналитики, включающей анализ больших данных, и улучшением показателей внутреннего отдела *Customer Relationship Management*.

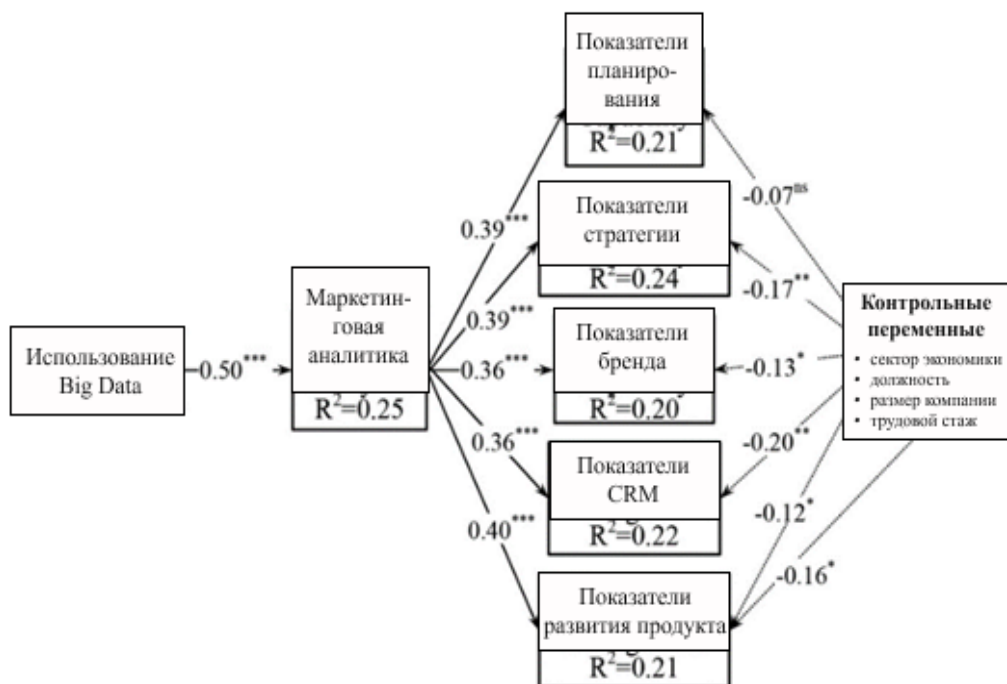


Рисунок 5. Результаты исследования взаимосвязи *Big Data* с результативностью китайских компаний [6]

Такая незначительная корреляция может быть обусловлена не столько низким влиянием *Big Data-driven* подхода на экономические показатели, сколько тем, что специалистам в области маркетинга не хватает понимания того, как использовать столь большие объемы информации для оптимизации бизнес-процессов. Не последнюю роль играют и региональные особенности: законодательство в области сбора персональных данных в Китае.

Для выявления более точной зависимости необходимо проводить дальнейшие исследования на рынке не одной страны, задействуя в опросах предприятия различных секторов экономики и размеров. Пробел в понимании технологии указывает и на то, что для небольших команд становится все более актуальным кросс-дисциплинарное обучение специалистов таким образом, чтобы те умели работать с большими объемами данных и принимать на их основе маркетинговые решения.

#### **Применение *Big Data* в сфере конкурентного анализа и анализа целевой аудитории.**

Для малых бизнесов, желающих сделать первые шаги в области аналитики данных с небольшими вложениями маркетингового бюджета и измеримой пользой, могут стать облачные инструменты веб-аналитики для конкурентного анализа.

Такие инструменты, как правило, предлагаются компаниями веб-аналитики, специализирующимися на веб-трафике и перформансе. Одним из примеров может послужить *SimilarWeb*: инструмент позволяет проводить мониторинг цифровых медиакампаний конкурентов с целью оптимизации собственных медиа-партнерств и поиска СМИ, рекламных сетей и потенциальных маркетинговых коллабораций. Основатели израильской компании *SimilarWeb Ltd.* рекомендуют использовать инструмент в сочетании с *Google Analytics*, подчеркивая их взаимодополняемость.

Распространенной альтернативой инструменту конкурентного анализа *SimilarWeb* является SaaS-платформа *SemRush*. Считается, что инструмент подходит в большей степени стартапам и малым бизнесам, в то время как *SimilarWeb* отвечает нуждам среднего бизнеса. Функционал платформы *SemRush* позволяет проводить следующие операции конкурентного анализа:

- разобрать маркетинговый микс конкурентов по их источникам трафика: прямым, поисковым, реферальным, платным и социальным;
- выяснить данные онлайн-рейтинга конкурентов: объем поиска, ключевые слова, стоимость за клик;
- узнать долю мобильного и десктопного трафика компаний на рынке;
- определить, из каких стран поступает трафик;
- изучить, какие сайты посещают люди до и после перехода на сайт конкурента.

Помимо облачных платформ, работающих по модели платной подписки, малый бизнес может воспользоваться бесплатными сервисами для анализа целевой аудитории и конкурентов от таких компаний, как Яндекс (так, Яндекс *Wordstat* позволяет узнать, какие запросы использует аудитория для поиска информации; Яндекс Метрика позволяет анализировать поведение целевой аудитории на сайте), Google (*Google Analytics* обладает аналогичным функционалом с Яндекс Метрикой; *Google Trends* позволяет анализировать сезонность и тренды, прогнозировать спрос на продукты и услуги компании), а также сервисами бизнес-модели «Фриминум» (*Freemium*), предоставляющей возможность воспользоваться базовым функционалом платформ бесплатно: так, действующий по *Freemium*-модели инструмент веб-аналитики *Serpstat* предоставляет следующие возможности:

- анализ обратных ссылок, отслеживание рейтинга и аудит сайта;
- выявление ботов по «всплескам» трафика из других стран;
- анализ контекстной рекламы конкурентов: сроки рекламных кампаний, ключевые слова, средняя стоимость клика, переходы по органическим словам.

#### **Безопасность данных в маркетинге.**

С расширением сферы использования данных о потребителях в маркетинге и особенно в цифровом маркетинге актуальным становится и вопрос безопасности данных.

Исследование *Global Data Protection Index*, проведенное *Dell EMC*, показало, что 81% предприятий обеспокоены рисками безопасности и конфиденциальности данных, связанными с облачными вычислениями и аналитикой больших данных [7].

Согласно исследованию *IBM*, потери компании в случае утечки данных составляют в среднем 3,86 миллиона долларов [8].

Обеспокоенность со стороны потребителей также является барьером на пути внедрения технологий *Big Data*.

Так, согласно исследованию, проведенному компанией *Accenture*, 73% потребителей готовы поделиться своими данными, если они уверены, что их данные используются прозрачно и этично.

В свою очередь более 51% опрошенных признают, что беспокоятся за обеспечение защиты сведений и их конфиденциальности (рисунок 4).

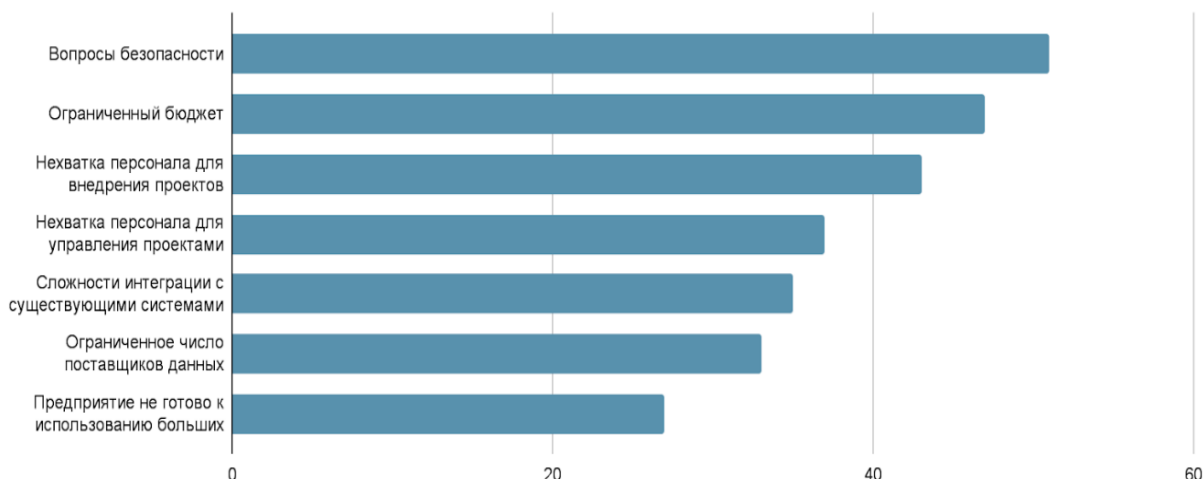


Рисунок 6. Основные проблемы при внедрении проектов *Big Data* [9]

Чтобы адресовать возрастающую обеспокоенность потребителей проблемами конфиденциальности, компаниям необходимо разработать политику безопасности данных, которая описывает, как предприятие собирает, использует и защищает персональные данные своих клиентов. Политика должна быть доведена до сведения сотрудников и клиентов и регулярно пересматриваться для обеспечения соответствия меняющимся нормативным требованиям и локальному законодательству.

Помимо этого, компаниям следует инвестировать в информационные технологии безопасности: брандмауэры, шифрование и системы обнаружения вторжений. При разработке маркетинговых кампаний предприятиям следует собирать только те данные, которые необходимы для их маркетинговой деятельности, и получать согласие клиентов перед сбором любых персональных данных.

Таким образом компании смогут укрепить доверие клиентов, а также повысить их лояльность и вовлеченность в цифровых каналах продвижения: веб-сайт, социальные медиа, реклама.

### Заключение.

Большие данные становятся важным инструментом для компаний, которые стремятся глубже понять своих клиентов и персонализировать свои предложения. Используя возможности аналитики *Big Data*, компании могут получить ценные сведения о моделях поведения и предпочтениях потребителей, удовлетворить их потребности и стимулировать рост бизнеса.

Однако важно, чтобы предприятия подходили к использованию больших данных этично и ответственно, принимая во внимание вопросы конфиденциальности и безопасности данных. Поскольку объем данных, генерируемых клиентами, продолжает расти, аналитика больших данных будет и дальше играть решающую роль в оказании помощи предприятиям.

### Список литературы

- [1] Алексеев К.А. Использование Big Data в международном бизнесе. Труды ИСП РАН, том 32, вып. 4, 2020 г., стр. 7-20.
- [2] Statista Report: Data Created Annually Worldwide from 2010 to 2025 [Электронный ресурс]. URL: <https://www.statista.com/statistics/871513/worldwide-data-created/> (Дата обращения: 25.03.2025).
- [3] IBM Institute for Business Value Report [Электронный ресурс]. URL: <https://www.ibm.com/thought-leadership/institute-business-value/en-us> (Дата обращения: 25.03.2025).
- [4] Wamba SF, Gunasekaran A, Akter S, Ren SJF, Dubey R, Childe SJ. Big data analytics and firm performance: effects of dynamic capabilities. *J Bus Res*.
- [5] Cao, Guangming & Tian, Na & Blankson, Charles. (2021). Big Data, Marketing Analytics, and Firm Marketing Capabilities. *Journal of Computer Information Systems*. 62. 1-10.
- [6] Anshari, Muhammad & Almunawar, Mohammad Nabil & Lim, Syamimi & Al-Mudimigh, Abdullah. (2018). Customer Relationship Management and Big Data Enabled: Personalization & Customization of Services. *Applied Computing and Informatics*.

[7] Dell Global Data Protection Index [Электронный ресурс]. URL: <https://www.delltechnologies.com/asset/en-nz/products/data-protection/industry-market/global-data-protection-index-key-findings.pdf> (Дата обращения: 25.03.2025).

[8] IBM Cost of a Data Breach Report 2021 [Электронный ресурс]. URL: <https://www.ibm.com/downloads/cas/OJDVQGRY#> (Дата обращения: 25.03.2025).

[9] Accenture Interactive 2019 Consumer Pulse Survey [Электронный ресурс]. URL: [https://www.accenture.com/\\_acnmedia/PDF-113/Accenture-Interactive-2019-Consumer-Pulse-Survey.pdf](https://www.accenture.com/_acnmedia/PDF-113/Accenture-Interactive-2019-Consumer-Pulse-Survey.pdf) (Дата обращения: 25.03.2025).

## **USING BIG DATA TO UNCOVER CONSUMER BEHAVIOR PATTERNS AND ADDRESS CUSTOMERS' NEEDS IN THE DIGITAL LANDSCAPE**

***E.S. Diagel***

*Student, Faculty of Engineering and Economics*

***D.A. Frolova***

*Lecturer, Department of Economics*

*Department of Economics*

*Faculty of Engineering and Economics*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: lizadziahel@gmail.com*

**Abstract.** The introduction of Big Data has revolutionized the way businesses understand their customers and interact with them. By means of an in-depth analysis of information gathered from various resources, Big Data enables marketers to gain valuable insights into customer behavior, as well as understand their preferences and needs better. However, it is important to take into account data privacy concerns, as the companies' ways of addressing data security issues play crucial role in their customers' loyalty and trust.

**Keywords:** Big Data, data-driven marketing, data analysis, business optimization, marketing analytics.

УДК 001.891

## РЕАЛИЗАЦИЯ РЕГИСТРАЦИИ УЧАСТНИКОВ НА БАЗЕ ЭЛЕКТРОННОЙ ПЛОЩАДКИ GOOGLE FORMS С ФОРМИРОВАНИЕМ ОНЛАЙН БАЗЫ ДАННЫХ



**В.Ф. Алексеев**

Доцент кафедры  
проектирования

информационно-компьютерных  
систем БГУИР, кандидат  
технических наук, доцент  
[alexvikt.minsk@gmail.com](mailto:alexvikt.minsk@gmail.com)



**Г.А. Пискун**

Доцент кафедры  
проектирования

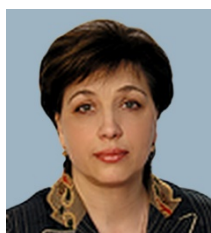
информационно-компьютерных  
систем БГУИР, кандидат  
технических наук, доцент  
[piskunbsuir@gmail.com](mailto:piskunbsuir@gmail.com)



**Д.В. Лихачевский**

Декан факультета

компьютерного проектирования  
БГУИР, кандидат технических  
наук, доцент  
[likhachevskiyd@bsuir.by](mailto:likhachevskiyd@bsuir.by)



**И.Н. Тонкович**

Доцент кафедры  
проектирования

информационно-компьютерных  
систем БГУИР, кандидат  
химических наук, доцент  
[intonkovich@gmail.com](mailto:intonkovich@gmail.com)



**Э.В. Асадчая**

Магистрант гр.115401,  
ассистент кафедры

проектирования  
информационно-компьютерных  
систем БГУИР  
[elina.asadchaya@gmail.com](mailto:elina.asadchaya@gmail.com)



**А.Д. Ларькин**

Магистрант гр.215441,  
ассистент кафедры

проектирования  
информационно-компьютерных  
систем БГУИР  
[anton11061998@gmail.com](mailto:anton11061998@gmail.com)

### **В.Ф. Алексеев**

Окончил Минский радиотехнический институт. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

### **Г.А. Пискун**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с моделированием и оптимальным проектированием информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

### **Д.В. Лихачевский**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием проблем радиочастотной идентификации объектов, организацией учебного и научно-исследовательского процессов в техническом университете.

### **И.Н. Тонкович**

Окончила Белорусский государственный университет. Область научных интересов связана с разработкой методов и алгоритмов компетентного подхода в подготовке специалистов, инновационного подхода в системе высшего образования, организацией учебного и научно-исследовательского процессов в техническом университете.

**Э.В. Асадчая**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой методов совершенствования научно-исследовательской работы студентов, организацией учебного и научно-исследовательского процессов в техническом университете.

**А.Д. Ларькин**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой методов и алгоритмов компьютерного инжиниринга, организацией учебного и научно-исследовательского процессов в техническом университете.

**Аннотация.** Представлена электронная *online*-форма регистрации участников конференции. Разработан скрипт для целевой отправки статей по секциям. Реализована эффективная регистрация участников с возможностью прикрепления и пересылки материала автора.

**Ключевые слова:** электронная *online*-форма регистрации, скрипт для целевой отправки, научная конференция.

**Введение.**

Научно-техническая работа в вузе предусматривает улучшение процесса подготовки высококвалифицированных специалистов, ведёт к повышению квалификации профессорско-преподавательского состава, позволяет вовлечь учёных, аспирантов и студентов в решении актуальных народно-хозяйственных проблем [1].

Участие аспирантов, магистрантов и студентов в научных конференциях предполагает апробацию результатов исследований, выполненных ими в течение определённого промежутка времени. Успех проведения любого научно-технического мероприятия зависит от правильности построения алгоритма действий по их организации. Это, в свою очередь, выдвигает определённые требования по организации электронной регистрации участников. Разработка электронных форм регистрации для проведения конференций повышает уровень достоверности информации об участниках и научных руководителях и выполнить статистическую обработку поступивших данных [1–8].

Современные требования к проведению конференций, приводят к необходимости использования таких новых форм и видов регистрации как, электронные формы участников. Рассмотрение данных форм началось относительно недавно, что подтверждается отсутствием данных по их созданию на примере конференций. Данные формы нашли наибольшее применение при проведении научных конференций в учреждениях образования (далее УО).

В связи с вышесказанным, актуальной является разработка методов совершенствования научно-исследовательской работы (далее НИР) студентов, участвующих в студенческой научной конференции [4–6].

**Реализация электронной формы регистрации.**

Для прохождения электронной формы регистрации необходимо наличие *Google* аккаунта.

1. Поле для ввода электронной почты для ведения переписки которое имеет несколько вариаций следующего вида:

- при открытии данной страницы имеет вид, представленный на рисунке 3.1 а;
- в случае, если данное поле не заполнено или заполнено некорректно, выдаются ошибки следующего содержания: «Это обязательный вопрос», «Укажите допустимый адрес электронной почты» и имеет вид, представленный на рисунке 3.1 б;
- корректно заполненное поле представлено на рисунке 3.1 в.



Регистрация участников 59-й научной конференции аспирантов, магистрантов и студентов УО «БГУИР»: направление "Электронные системы и технологии"

Даты проведения: 17-21 апреля 2023 г.  
Даты приема заявок: 16 января - 24 марта 2023 г.

elina.asadchaya@gmail.com [Сменить аккаунт](#)

Когда вы загрузите файлы и отправите форму, мы сохраним ваши имя и фото профиля. В ответе будет использован только введенный вами адрес электронной почты.

\* Обязательно

Электронная почта \*

Ваш адрес эл. почты

elina.asadchaya@gmail.com

Анкета участника  
Для участия в конференции, пожалуйста, заполните приведенную ниже форму и нажмите кнопку "Отправить"

Ссылки на документы:  
1. Информационное письмо конференции: <https://drive.google.com/file>  
2. Шаблон оформленной статьи: <https://docs.google.com/document>  
3. Сокращение ученых степеней и званий: <https://drive.google.com/file>  
4. Правила оформления использованных источников: <https://drive.google.com/file>

Далее Очистить форму

а)

б)

в)

Рисунок 1. Реализация первой страницы регистрации

2. Информация о том, что за форму регистрации заполняют участники и что необходимо сделать: «Анкета участника. Для участия в конференции, пожалуйста, заполните приведенную ниже форму и нажмите кнопку "Отправить"».

3. Дополнительная информация для участников конференции, содержит следующие документы: информационное письмо, включающее в себя основные положения по приему материалов и проведению конференции; шаблон оформленной статьи; сокращение ученых степеней и званий; правила оформления использованных источников.

**Первый раздел** электронной формы регистрации необходим для сбора информации об авторе (в описании раздела можно сразу заметить – максимальное количество авторов – 3), данный раздел представлен на рисунке 2 а, б, в и содержит следующую информацию:

1. Поле «Автор». Данное поле предназначено для ввода фамилии, имени и отчества автора.

2. Поле «Укажите страну» предназначено для ввода информации о стране автора.

3. Поле «Укажите город» предназначено для ввода информации о городе автора.

4. Поле «Укажите учреждение образования» предназначено для ввода информации о учреждении образования автора.

5. Поле «Укажите кафедру, на которой обучаетесь» предназначено для ввода информации о кафедре, на которой обучается автор.

6. Поле «Степень обучения автора» данный вопрос был реализован в виде выпадающего списка с вариативностью ступеней, в зависимости от выбора ступени следующий вопрос будет либо обязательным (студент, магистрант, курсант), либо нет (аспирант).

7. Поле «Номер группы автора» данное поле не является обязательным в связи с тем, что в данной конференции принимаю участие различные градации обучающихся.

8. Поле «Электронная почта автора» предназначено для ввода дополнительной электронной почты.

Поля, указанные в пунктах 5–10, 12, являются обязательными и имеют несколько вариаций следующего вида:

- при открытии данной страницы имеет вид, представленный на рисунке 2 а;
- в случае если данное поле не заполнено или заполнено некорректно выдаются ошибки следующего содержания: «Это обязательный вопрос», «Проверьте правильность написания данных» и имеет вид, представленный на рисунке 2 б;
- корректно заполненное поле представлено на рисунке 2 в.

В связи с тем, что научных руководителей может быть несколько, были **реализованы два раздела (второй, третий)** для сбора информации о научных руководителях (рисунок 3 а, б, в – 4 а, б, в), которые включают в себя следующие поля для ввода данных:

1. Поле «Научный руководитель». Данное поле предназначено для ввода фамилии, имени и отчества научного руководителя.

2. Поле «Ученая степень» предназначено для ввода информации об ученой степени научного руководителя.

3. Поле «Ученое звание» предназначено для ввода информации об ученом звании научного руководителя.

4. Поле «Должность научного руководителя» предназначено для ввода информации о должности научного руководителя.

5. Поле «Второй научный руководитель». Данное поле было реализовано в виде списка с возможностью выбора и предназначено для ввода информации о наличии или отсутствии второго научного руководителя. В зависимости от ответа переход осуществляется либо к разделу про второго научного руководителя (третий раздел), либо к четвертому разделу. Данный вопрос присутствует только в этом разделе (второй раздел).

Все поля, указанные в данных разделах, являются обязательными и имеют несколько вариаций следующего вида:

- при открытии данной страницы имеет вид, представленный на рисунке 3 а, 4 а;
- в случае если данное поле не заполнено или заполнено некорректно выдаются ошибки следующего содержания: «Это обязательный вопрос», «Проверьте правильность написания данных» и имеет вид, представленный на рисунке 3 б, 4 б;
- корректно заполненное поле представлено на рисунке 3 в, 4 в.

Информация об авторе	Информация об авторе	Информация об авторе
Максимальное количество авторов – 3 человека	Максимальное количество авторов – 3 человека	Максимальное количество авторов – 3 человека
<p>Автор *</p> <p>Например: Иванов Иван Иванович</p> <p>Мой ответ _____</p>	<p>Автор *</p> <p>Например: Иванов Иван Иванович</p> <p>Мой ответ _____</p> <p><b>!</b> Это обязательный вопрос.</p>	<p>Автор *</p> <p>Например: Иванов Иван Иванович</p> <p>Асадчая Элина Вячеславовна _____</p>
<p>Укажите страну *</p> <p>Например: Республика Беларусь</p> <p>Мой ответ _____</p>	<p>Укажите страну *</p> <p>Например: Республика Беларусь</p> <p>Мой ответ _____</p> <p><b>!</b> Это обязательный вопрос.</p>	<p>Укажите страну *</p> <p>Например: Республика Беларусь</p> <p>Республика Беларусь _____</p>
<p>Укажите город *</p> <p>Например: Минск</p> <p>Мой ответ _____</p>	<p>Укажите город *</p> <p>Например: Минск</p> <p>Мой ответ _____</p> <p><b>!</b> Это обязательный вопрос.</p>	<p>Укажите город *</p> <p>Например: Минск</p> <p>Минск _____</p>
<p>Укажите учреждение образования *</p> <p>Например: Белорусский государственный университет информатики и радиоэлектроники (УО "БГУИР") Московский государственный лингвистический университет (ФГБОУ ВО МГЛУ)</p> <p>Мой ответ _____</p>	<p>Укажите учреждение образования *</p> <p>Например: Белорусский государственный университет информатики и радиоэлектроники (УО "БГУИР") Московский государственный лингвистический университет (ФГБОУ ВО МГЛУ)</p> <p>Мой ответ _____</p> <p><b>!</b> Это обязательный вопрос.</p>	<p>Укажите учреждение образования *</p> <p>Например: Белорусский государственный университет информатики и радиоэлектроники (УО "БГУИР") Московский государственный лингвистический университет (ФГБОУ ВО МГЛУ)</p> <p>Белорусский государственный университет информатики и радиоэлектроники (УО «БГУИР») _____</p>
<p>Укажите кафедру на которой обучаетесь *</p> <p><input type="radio"/> "Инженерная и компьютерная графика" (каф. ИКГ)</p> <p><input type="radio"/> "Инженерная психология и эргономика" (каф. ИПиЭ)</p> <p><input type="radio"/> "Иностранные языки" (каф. ИнЯз)</p> <p><input type="radio"/> "Проектирование информационно-компьютерных систем" (каф.ПИКС)</p> <p><input type="radio"/> "Электронная техника и технологии" (каф. ЭТТ)</p> <p><input type="radio"/> Другое: _____</p>	<p>Укажите кафедру на которой обучаетесь *</p> <p><input type="radio"/> "Инженерная и компьютерная графика" (каф. ИКГ)</p> <p><input type="radio"/> "Инженерная психология и эргономика" (каф. ИПиЭ)</p> <p><input type="radio"/> "Иностранные языки" (каф. ИнЯз)</p> <p><input type="radio"/> "Проектирование информационно-компьютерных систем" (каф.ПИКС)</p> <p><input type="radio"/> "Электронная техника и технологии" (каф. ЭТТ)</p> <p><input type="radio"/> Другое: _____</p>	<p>Укажите кафедру на которой обучаетесь *</p> <p><input type="radio"/> "Инженерная и компьютерная графика" (каф. ИКГ)</p> <p><input type="radio"/> "Инженерная психология и эргономика" (каф. ИПиЭ)</p> <p><input type="radio"/> "Иностранные языки" (каф. ИнЯз)</p> <p><input checked="" type="radio"/> "Проектирование информационно-компьютерных систем" (каф.ПИКС)</p> <p><input type="radio"/> "Электронная техника и технологии" (каф. ЭТТ)</p> <p><input type="radio"/> Другое: _____</p>
<p>Степень обучения автора *</p> <p>Выбрать ▾</p>	<p>Степень обучения автора *</p> <p>Выбрать ▾</p> <p><b>!</b> Это обязательный вопрос.</p>	<p>Степень обучения автора *</p> <p>Магистрант ▾</p>
<p>Номер группы автора (ТОЛЬКО для студентов и магистрантов)</p> <p>Мой ответ _____</p>	<p>Номер группы автора (ТОЛЬКО для студентов и магистрантов)</p> <p>Мой ответ _____</p>	<p>Номер группы автора (ТОЛЬКО для студентов и магистрантов)</p> <p>115401 _____</p>
<p>Электронная почта автора *</p> <p>Например: <a href="mailto:Abc123@gmail.com">Abc123@gmail.com</a></p> <p>Мой ответ _____</p>	<p>Электронная почта автора *</p> <p>Например: <a href="mailto:Abc123@gmail.com">Abc123@gmail.com</a></p> <p>Мой ответ _____</p> <p><b>!</b> Это обязательный вопрос.</p>	<p>Электронная почта автора *</p> <p>Например: <a href="mailto:Abc123@gmail.com">Abc123@gmail.com</a></p> <p><a href="mailto:elina.asadchaya@gmail.com">elina.asadchaya@gmail.com</a> _____</p>

а)

б)

в)

Рисунок 2. Фрагмент реализации раздела «Информация об авторе» в электронной форме регистрации

Информация о научном руководителе

Научный руководитель \*  
Например: Иванов Иван Иванович  
Мой ответ

Ученая степень \*  
Например: кандидат технических наук  
Мой ответ

Ученое звание \*  
Например: доцент  
Мой ответ

Должность научного руководителя \*  
Например: доцент кафедры ПИКС  
Мой ответ

Второй научный руководитель \*  
 Да  
 Нет

Информация о научном руководителе

Научный руководитель \*  
Например: Иванов Иван Иванович  
Мой ответ  
❗ Это обязательный вопрос.

Ученая степень \*  
Например: кандидат технических наук  
Мой ответ  
❗ Это обязательный вопрос.

Ученое звание \*  
Например: доцент  
Мой ответ  
❗ Это обязательный вопрос.

Должность научного руководителя \*  
Например: доцент кафедры ПИКС  
Мой ответ  
❗ Это обязательный вопрос.

Второй научный руководитель \*  
 Да  
 Нет

Информация о научном руководителе

Научный руководитель \*  
Например: Иванов Иван Иванович  
Пискун Геннадий Адамович

Ученая степень \*  
Например: кандидат технических наук  
Кандидат технических наук

Ученое звание \*  
Например: доцент  
Доцент

Должность научного руководителя \*  
Например: доцент кафедры ПИКС  
Доцент кафедры ПИКС

Второй научный руководитель \*  
 Да  
 Нет

a) b) в)

Рисунок 3. Фрагмент реализации раздела «Информация о научном руководителе» электронной формы регистрации

Информация о научном руководителе

Научный руководитель \*  
Например: Иванов Иван Иванович  
Мой ответ

Ученая степень \*  
Например: кандидат технических наук  
Мой ответ

Ученое звание \*  
Например: доцент  
Мой ответ

Должность научного руководителя \*  
Например: доцент кафедры ПИКС  
Мой ответ

Второй научный руководитель \*  
 Да  
 Нет

Информация о научном руководителе

Научный руководитель \*  
Например: Иванов Иван Иванович  
Мой ответ  
❗ Это обязательный вопрос.

Ученая степень \*  
Например: кандидат технических наук  
Мой ответ  
❗ Это обязательный вопрос.

Ученое звание \*  
Например: доцент  
Мой ответ  
❗ Это обязательный вопрос.

Должность научного руководителя \*  
Например: доцент кафедры ПИКС  
Мой ответ  
❗ Это обязательный вопрос.

Второй научный руководитель \*  
 Да  
 Нет

Информация о научном руководителе

Научный руководитель \*  
Например: Иванов Иван Иванович  
Алексеев Виктор Федорович

Ученая степень \*  
Например: кандидат технических наук  
Кандидат технических наук

Ученое звание \*  
Например: доцент  
Доцент

Должность научного руководителя \*  
Например: доцент кафедры ПИКС  
Доцент кафедры ПИКС

Второй научный руководитель \*  
 Да  
 Нет

a) b) в)

Рисунок 4. Фрагмент реализации раздела «Информация о втором научном руководителе» электронной формы регистрации

**Четвертый раздел** электронной формы регистрации необходим для сбора информации о публикации, данный раздел представлен на рисунке 5 а, б, в и содержит следующую информацию:

1. Поле «Название статьи (утвержденное научным руководителем)». Данное поле предназначено для ввода название статьи, примечание указано для исключения ситуаций с получением статей непроверенных научным руководителем.

2. Поле «Форма участия – дистанционная». Данное поле было реализовано в виде списка с возможностью выбора (выступление, публикация либо выступление и публикация):

- публикация работы и выступление с докладом (с выдачей электронного сертификата);
- публикация работы без выступления с докладом;
- выступление с докладом.

3. Поле «Пишете в соавторстве». Данное поле также было реализовано в виде списка с возможностью выбора и предназначено для ввода информации о наличии или отсутствии первого соавтора. В зависимости от ответа переход осуществляется либо к разделу про первого соавтора (пятый раздел), либо к седьмому разделу.

Все поля, указанные в данном разделе, являются обязательными и имеют несколько вариаций следующего вида:

- при открытии данной страницы имеет вид, представленный на рисунке 5а;
- в случае если данное поле не заполнено или заполнено некорректно выдаются ошибки следующего содержания: «Это обязательный вопрос», «Проверьте правильность написания данных» и имеет вид, представленный на рисунке 5 б;
- корректно заполненное поле представлено на рисунке 5 в.

The figure shows three variations of a web form titled 'Информация о публикации' (Information about publication). Each variation has three main sections:

- Section 1:** 'Название статьи (утвержденной научным руководителем)' (Article title) with a red asterisk and a 'Мой ответ' (My answer) input field.
- Section 2:** 'Форма участия - дистанционная' (Participation form) with three radio button options: 'Публикация работы и выступление с докладом (с выдачей электронного сертификата)', 'Публикация работы без выступления с докладом', and 'Выступление с докладом'.
- Section 3:** 'Пишете в соавторстве?' (Co-authorship?) with two radio button options: 'Нет' (No) and 'Да' (Yes).

Variant **а)** shows the form in its initial state with all fields empty.

Variant **б)** shows the form with red error messages: 'Это обязательный вопрос.' (This is a mandatory question.) next to the empty title field and the empty participation form.

Variant **в)** shows the form with the participation form filled out: the second option 'Публикация работы без выступления с докладом' is selected with a blue dot.

Рисунок 5. Фрагмент реализации раздела «Информация о публикации» электронной формы регистрации

В связи с тем, что соавторов может быть несколько, были реализованы два раздела (пятый, шестой) для сбора информации о соавторах (рисунок 6 а, б, в – 7 а, б, в), которые включают в себя следующие поля для ввода данных:

1. Поле «Соавтор». Данное поле предназначено для ввода фамилии, имени и отчества соавтора.

2. Поле «Степень обучения соавтора» предназначено для ввода информации об ученой степени научного руководителя. Данный вопрос был реализован в виде выпадающего списка с вариативностью ступеней, в зависимости от выбора ступени следующий вопрос будет либо обязательным (студент, магистрант, курсант), либо нет (аспирант).

3. Поле «Номер группы соавтора». Данное поле не является обязательным в связи с тем, что в данной конференции принимаю участие различные градации обучающихся.

4. Поле «Следующий соавтор». Данное поле было реализовано в виде списка с возможностью выбора и предназначено для ввода информации о наличии или отсутствии второго соавтора. В зависимости от ответа переход осуществляется либо к разделу про второго соавтора (шестой раздел), либо к седьмому разделу. Данный вопрос присутствует только в этом разделе (пятый раздел).

В данном разделе поля 1–2, 4 являются обязательными и имеют несколько вариаций следующего вида:

- при открытии данной страницы имеет вид, представленный на рисунке 6 а, 7 а;
- в случае если данное поле не заполнено или заполнено некорректно выдаются ошибки следующего содержания: «Это обязательный вопрос», «Проверьте правильность написания данных» и имеет вид, представленный на рисунке 6 б, 7 б;
- корректно заполненное поле представлено на рисунке 6 в, 7 в.

а) **Информация о соавторе**  
Соавтор \*  
Например: Иванов Иван Иванович  
Мой ответ  
Степень обучения соавтора \*  
Выбрать  
Номер группы соавтора (ТОЛЬКО для студентов и магистрантов)  
Мой ответ  
Следующий соавтор \*  
После заполнения нажмите на кнопку "Далее"  
 Нет  
 Да

б) **Информация о соавторе**  
Соавтор \*  
Например: Иванов Иван Иванович  
Мой ответ  
**!** Это обязательный вопрос.  
Степень обучения соавтора \*  
Выбрать  
**!** Это обязательный вопрос.  
Номер группы соавтора (ТОЛЬКО для студентов и магистрантов)  
Мой ответ  
Следующий соавтор \*  
После заполнения нажмите на кнопку "Далее"  
 Нет  
 Да

в) **Информация о соавторе**  
Соавтор \*  
Например: Иванов Иван Иванович  
Калита Елена Викторовна  
Степень обучения соавтора \*  
Магистрант  
Номер группы соавтора (ТОЛЬКО для студентов и магистрантов)  
115401  
Следующий соавтор \*  
После заполнения нажмите на кнопку "Далее"  
 Нет  
 Да

Рисунок 6. Фрагмент реализации раздела «Информация о соавторе» электронной формы регистрации



а)

б)

в)

Рисунок 7. Фрагмент реализации раздела «Информация о втором соавторе» электронной формы регистрации

**Седьмой раздел** электронной формы регистрации необходим для сбора информации о выборе секции, данный раздел представлен на рисунке 8 а, б, в и содержит следующую информацию:

1. Ссылки для предоставления материалов конференции (статей). Данное поле включает в себя подробную информацию о секциях и их направлениях конференции:

2. Поле «В какой секции планируете участвовать?». Данное поле было реализовано в виде списка с возможностью выбора. В зависимости от ответа прикрепленный в последнем восьмом разделе файл со статьей будет отправляться на почту секции в соответствии с выбранной позицией. Данное поле является обязательными и имеют несколько вариаций следующего вида:

- при открытии данной страницы имеет вид, представленный на рисунке 8 а;
- в случае если данное поле не заполнено выдается ошибка следующего содержания: «Это обязательный вопрос», и имеет вид, представленный на рисунке 8 б;
- корректно заполненное поле представлено на рисунке 8 в.

**Восьмой раздел** электронной формы регистрации необходим для сбора информации о хранении и отправке данных, данный раздел представлен на рисунке 9 а, б, в и содержит следующую информацию:

1. Поле «Файл» предназначено для прикрепления файла с готовой оформленной статьей. Допустимы форматы файла: .doc, .docx. Необходимость такого формата заключается в том, что в последующих этапах при верстке сборника используются только данные форматы.

2. Поле «Согласие на хранение персональных данных». Данное поле было реализовано в виде списка с возможностью выбора. В зависимости от ответа переход осуществляется либо к отправке данных, либо к начальной странице. Также у этого поля имеется примечание:

*Примечание: Я согласен на обработку Оператором своих персональных данных в соответствии с требованиями Закона Республики Беларусь от 07.05.2021 № 99-3 "О защите персональных данных". После ответа на вопрос, нажмите кнопку "Отправить".*

Данное поле является обязательными и имеют несколько вариаций следующего вида:

- при открытии данной страницы имеет вид, представленный на рисунке 9 а;
- в случае если данное поле не заполнено выдается ошибка следующего содержания: «Это обязательный вопрос», и имеет вид, представленный на рисунке 9 б;
- корректно заполненное поле представлено на рисунке 9 в.

Выбор секции	Выбор секции	Выбор секции
<p>Ссылки для предоставления материалов конференции (статей): *Студентам, магистрантам и аспирантам иностранных и иногородних вузов, просьба выбрать секцию исходя из направления вашего исследования</p> <p><u>Секция «Проектирование информационно-компьютерных систем»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Моделирование и компьютерное проектирование радиоэлектронных средств»</li> <li>«Электронные системы безопасности»</li> <li>«Информационные системы и технологии в бизнес-менеджменте»</li> <li>«Программируемые мобильные системы»</li> </ul> <p><u>Секция «Электронная техника и технологии»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Программно-управляемые электронно-оптические системы»</li> <li>«Проектирование и производство программно-управляемых электронных средств»</li> <li>«Медицинская электроника»</li> </ul> <p><u>Секция «Инженерная психология, когнитивная эргономика, экологическая и промышленная безопасность»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Инженерно-психологическое обеспечение информационных технологий»</li> <li>«Информационные системы и технологии в обеспечении промышленной безопасности»</li> </ul> <p><u>Секция «Инженерная и компьютерная графика»</u></p> <p><u>Секция «Иностранные языки»</u></p>	<p>Ссылки для предоставления материалов конференции (статей): *Студентам, магистрантам и аспирантам иностранных и иногородних вузов, просьба выбрать секцию исходя из направления вашего исследования</p> <p><u>Секция «Проектирование информационно-компьютерных систем»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Моделирование и компьютерное проектирование радиоэлектронных средств»</li> <li>«Электронные системы безопасности»</li> <li>«Информационные системы и технологии в бизнес-менеджменте»</li> <li>«Программируемые мобильные системы»</li> </ul> <p><u>Секция «Электронная техника и технологии»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Программно-управляемые электронно-оптические системы»</li> <li>«Проектирование и производство программно-управляемых электронных средств»</li> <li>«Медицинская электроника»</li> </ul> <p><u>Секция «Инженерная психология, когнитивная эргономика, экологическая и промышленная безопасность»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Инженерно-психологическое обеспечение информационных технологий»</li> <li>«Информационные системы и технологии в обеспечении промышленной безопасности»</li> </ul> <p><u>Секция «Инженерная и компьютерная графика»</u></p> <p><u>Секция «Иностранные языки»</u></p>	<p>Ссылки для предоставления материалов конференции (статей): *Студентам, магистрантам и аспирантам иностранных и иногородних вузов, просьба выбрать секцию исходя из направления вашего исследования</p> <p><u>Секция «Проектирование информационно-компьютерных систем»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Моделирование и компьютерное проектирование радиоэлектронных средств»</li> <li>«Электронные системы безопасности»</li> <li>«Информационные системы и технологии в бизнес-менеджменте»</li> <li>«Программируемые мобильные системы»</li> </ul> <p><u>Секция «Электронная техника и технологии»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Программно-управляемые электронно-оптические системы»</li> <li>«Проектирование и производство программно-управляемых электронных средств»</li> <li>«Медицинская электроника»</li> </ul> <p><u>Секция «Инженерная психология, когнитивная эргономика, экологическая и промышленная безопасность»</u></p> <p>Направления:</p> <ul style="list-style-type: none"> <li>«Инженерно-психологическое обеспечение информационных технологий»</li> <li>«Информационные системы и технологии в обеспечении промышленной безопасности»</li> </ul> <p><u>Секция «Инженерная и компьютерная графика»</u></p> <p><u>Секция «Иностранные языки»</u></p>
<p>В какой секции планируете участвовать? *</p> <p><input type="radio"/> "Инженерная и компьютерная графика" (каф. ИКГ)</p> <p><input type="radio"/> "Инженерная психология, когнитивная эргономика, экологическая и промышленная безопасность" (каф. ИПиЭ)</p> <p><input type="radio"/> "Иностранные языки" (каф. ИнЯз)</p> <p><input type="radio"/> "Проектирование информационно-компьютерных систем" (каф.ПИКС)</p> <p><input type="radio"/> "Электронная техника и технологии" (каф. ЭТТ)</p>	<p>В какой секции планируете участвовать? *</p> <p><input type="radio"/> "Инженерная и компьютерная графика" (каф. ИКГ)</p> <p><input type="radio"/> "Инженерная психология, когнитивная эргономика, экологическая и промышленная безопасность" (каф. ИПиЭ)</p> <p><input type="radio"/> "Иностранные языки" (каф. ИнЯз)</p> <p><input type="radio"/> "Проектирование информационно-компьютерных систем" (каф.ПИКС)</p> <p><input type="radio"/> "Электронная техника и технологии" (каф. ЭТТ)</p> <p><input checked="" type="radio"/> Это обязательный вопрос.</p>	<p>В какой секции планируете участвовать? *</p> <p><input type="radio"/> "Инженерная и компьютерная графика" (каф. ИКГ)</p> <p><input type="radio"/> "Инженерная психология, когнитивная эргономика, экологическая и промышленная безопасность" (каф. ИПиЭ)</p> <p><input type="radio"/> "Иностранные языки" (каф. ИнЯз)</p> <p><input checked="" type="radio"/> "Проектирование информационно-компьютерных систем" (каф.ПИКС)</p> <p><input type="radio"/> "Электронная техника и технологии" (каф. ЭТТ)</p>

Рисунок 8. Фрагмент реализации раздела «Информация о секции» электронной формы регистрации

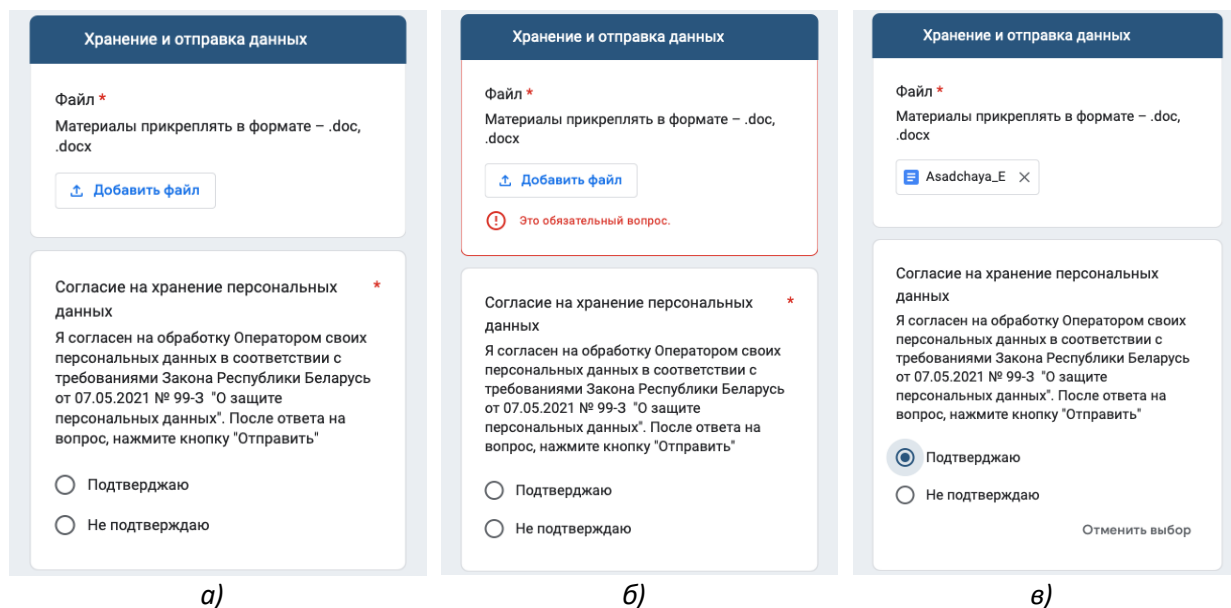


Рисунок 9. Фрагмент реализации раздела «Хранение и отправка данных» электронной формы регистрации

Для максимально удобного визуального оформления, практически у каждого вопроса есть описание, в котором указан пример правильного заполнения поля регистрации.

В каждом разделе также имеются обязательные поля, данная информация является основной для сбора статистических данных.

При заполнении всех обязательных полей каждого раздела происходит отправка данных в БД, а также для автоматизации электронной регистрации была оптимизирована форма, теперь при отправке формы прикрепленный файл научной работы отправляется на почту выбранной секции.

#### Заключение.

Представлена электронная *online*-форма регистрации участников конференции. Основной особенностью которой является наличие авторского скрипта для целевой отправки статей по секциям. Реализована эффективная регистрация участников с возможностью прикрепления и пересылки материала автора.

#### Список литературы

- [1] Лукьянец, С. В. Вне программы, в помощь программе / С. В. Лукьянец, В. Ф. Алексеев // Промышленность Белоруссии. – 1979. – №5. – С.48-49.
- [2] Алексеев, В. Ф. Методология организации научно-исследовательской работы студентов, обучающихся по дистанционной форме образования / В. Ф. Алексеев, Д. В. Лихачевский, Г. А. Пискун // Дистанционное обучение – образовательная среда XXI века : материалы X международной научно-методической конференции (Минск, 7 - 8 декабря 2017 года). – Минск : БГУИР, 2017. – С. 46 - 47.
- [3] Пискун, Г. А. Особенности применения активных методов обучения при подготовке специалистов в техническом вузе / А. Г. Пискун, В. Ф. Алексеев // Дистанционное обучение – образовательная среда XXI века : материалы IX международной научно-методической конференции (Минск, 3-4 декабря 2015 года). – Минск : БГУИР, 2015. – С. 83 – 84.
- [4] Пискун, Г.А. Оптимизация процесса проведения лабораторных занятий по дисциплине «Физические основы проектирования радиоэлектронных средств» / Г.А. Пискун, В.Ф. Алексеев / Дистанционное обучение – образовательная среда XXI века: материалы IX междунар. науч.-метод. конф. (Минск, 3-4 декабря 2015 г.). – Минск: БГУИР, 2015. – С.136-138.
- [5] Алексеев, В.Ф. Инженерное творчество в системе многоуровневого университетского образования / В.Ф. Алексеев, Д.В. Лихачевский, Г.А. Пискун // Дистанционное обучение – образовательная среда XXI века : материалы X международной научно-методической конференции (Минск, 7 – 8 декабря 2017 года). – Минск : БГУИР, 2017. – С. 124 – 125.

[6] Алексеев, В.Ф. Особенности обучения студентов в online формате / В.Ф. Алексеев, Г.А. Пискун, Д.В. Лихачевский // Дистанционное обучение – образовательная среда XXI века : материалы X международной научно-методической конференции (Минск, 7 – 8 декабря 2017 года). – Минск : БГУИР, 2017. – С. 136 – 137.

[7] Алексеев, В. Ф. Подходы к формированию университетской концепции развития научно-исследовательской работы аспирантов, магистрантов, и студентов в современных условиях / В. Ф. Алексеев, Л. С. Алексеева // Перспективы развития системы научно-исследовательской работы студентов в Республике Беларусь: сб. материалов науч.-практ. конф. / редкол. : А. И. Жук (пред.) [и др.]. – Минск: Изд. центр БГУ, 2011. – С. 29-38.

[8] Алексеев, В. Ф. Роль профилирующей кафедры в подготовке высококвалифицированных специалистов / В. Ф. Алексеев // Опыт и проблемы организации научно-исследовательской работы студентов: Сборник научных статей. – Мн.: БГУИР, 2003. – С. 18-21.

## **IMPLEMENTATION OF REGISTRATION OF PARTICIPANTS ON THE BASIS OF THE GOOGLE FORMS ELECTRONIC PLATFORM WITH THE FORMATION OF ONLINE DATABASE**

***V.F. Alexseev***

*Associate Professor of the  
Department of Design of  
Information and Computer Systems  
of BSUIR, PhD of Technical  
Sciences, Associate Professor*

***G.A. Piskun***

*Associate Professor of the  
Department of Design of  
Information and Computer  
Systems of BSUIR, PhD of  
Technical Sciences, Associate  
Professor*

***D.V. Likhachevsky***

*Dean of the Faculty of Computer  
Design of BSUIR,  
PhD of Technical Sciences,  
Associate Professor*

***I.N. Tonkovich***

*Associate Professor of the  
Department of Design of  
Information and Computer Systems  
of BSUIR, PhD of Technical  
Sciences, Associate Professor*

***E.V. Asadchaya***

*Master's student, Assistant of the  
Department of Design of  
Information and Computer  
Systems of BSUIR*

***A.D. Larkin***

*Master's student, Assistant of the  
Department of Design of  
Information and Computer Systems  
of BSUIR*

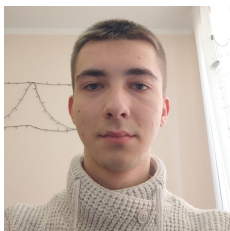
*Department of Information and Computer Systems Design  
Faculty of Computer Engineering  
Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: alexvikt.minsk@gmail.com*

**Abstract.** An electronic online form for registration of conference participants is presented. A script has been developed for targeted submission of articles in sections. Effective registration of participants with the possibility of attaching and forwarding the author's material has been implemented.

**Keywords:** electronic online registration form, script for targeted submission, scientific conference.

УДК 004.424.5

## ОПТИМИЗАЦИЯ ОБРАБОТКИ BIG DATA С ПОМОЩЬЮ ГРАФИЧЕСКИХ ПРОЦЕССОРОВ



**С.А. Байчик**

Студент 4-ого курса БГУИР  
ФКП, по специальности  
Проектирование и производство  
программно-управляемых  
электронных средств  
[s.bajchik@bsuir.by](mailto:s.bajchik@bsuir.by)



**С.Н. Нестеренков**

Кандидат технических наук,  
доцент, декан факультета  
компьютерных систем и сетей  
БГУИР, доцент кафедры ПОИТ  
[s.nesterenkov@bsuir.by](mailto:s.nesterenkov@bsuir.by)



**И.С. Тарасюк**

Инженер-программист ОИТ,  
ассистент кафедры ЭВМ  
[i.tarasiuk@bsuir.by](mailto:i.tarasiuk@bsuir.by)

### **С.А. Байчик**

Студент 4-ого курса БГУИР ФКП, по специальности Проектирование и производство программно-управляемых электронных средств.

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

### **И.С. Тарасюк**

Окончил БГУИР в 2021 году по специальности «Вычислительные машины, системы и сети», магистрант первого года обучения по специальности «Системы и сети инфокоммуникаций» БГУИР.

**Аннотация.** Данная статья посвящена оптимизации обработки больших объемов данных с использованием графических процессоров (GPU). В статье будет представлено общее представление о технологии GPU, ее особенностях и преимуществах при обработке больших данных. Основным объектом исследования статьи является анализ производительности и оптимизации обработки больших данных на GPU с использованием различных методов и техник, таких как параллельное программирование, оптимизация ядер и алгоритмов, использование специальных библиотек. Целью данной работы является представление и анализ различных подходов к оптимизации обработки больших данных на GPU, а также оценка эффективности этих подходов и их применимости в реальных условиях.

**Ключевые слова:** Big Data, графический процессор, производительность.

### **Введение.**

Уже несколько лет дата-центры многих компаний работают с вычислениями с ускорением на GPU (Graphics Processing Unit). Такой тип вычислений становится все более востребованным. Так, вычисления с ускорением на GPU можно (и нужно) использовать для ускорения требовательных к ресурсам приложений, созданных для работы в таких сферах, как глубокое обучение, аналитика и проектирование. Этот метод используют в дата-центрах крупных компаний, в лабораториях научно-исследовательских организаций, на предприятиях. Благодаря ускорению на GPU работают многие сервисы, обеспечивающие работу нейронных сетей или обрабатывающие данные, поступающие с умных автомобилей. Плюсом такого метода является то, что ресурсоемкая часть приложения, которая требует высокой вычислительной мощности, обрабатывается на GPU, а все остальное выполняется на центральном процессоре (CPU). В последние несколько лет стали появляться комбинированные решения, на основе которых создаются высокоскоростные базы данных. Такие решения стоит использовать, например, для визуализации крупных массивов данных.

### **Архитектура и особенности работы с GPU.**

Современные графические процессоры используют конвейерную архитектуру (graphic pipeline) для выполнения операций с максимальной эффективностью компьютерной графики. Эта структура позволяет достичь высокой производительности при работе с большими объемами данных благодаря потоковой параллельности.

Входные данные для GPU представляются набором вершин, которые поступают из приложения. Затем вершины обрабатываются в вершинных процессорах, которые классифицируются как MIMD. Для работы с вершинными процессорами используется программа, называемая вершинным шейдером. После обработки результат работы вершинного шейдера поступает на сборку примитивов, таких как полигоны или линии. Затем выполняются стандартные операции компьютерной графики, такие как тесты видимости отсечения. После этого данные поступают на растеризацию, где объемное изображение проецируется на плоский экран и масштабируется согласно параметрам окна приложения. Результатом этой операции является текстура, представляющая из себя прямоугольный массив пикселей. Пиксель может быть представлен 1-4 числами, например, в формате RGBA, т.е. red, green, blue, alpha (коэффициент прозрачности).

Текстура обрабатывается в процессорах пикселей, которые, согласно классификации многопроцессорных систем, могут быть отнесены к категории SIMD. После выполнения дополнительных операций данные передаются в буфер кадра, который и является изображением, видимым на экране для потребителя.

Архитектура GPU отличается от архитектуры CPU тем, что она специализирована на выполнении параллельных вычислений. GPU состоит из большого числа ядер (обычно от нескольких сотен до нескольких тысяч), которые могут работать независимо друг от друга. Это позволяет GPU эффективно обрабатывать множество задач одновременно.

GPU, аналогично CPU, обладает своими регистрами и кэшами для быстрого доступа к данным в процессе вычислений. В дополнение к этому у GPU есть своя основная память - графическая память. Для того, чтобы программист мог выполнять вычисления на GPU, ему необходимо предварительно передать данные из оперативной памяти CPU в память GPU. Эта операция традиционно считается дорогостоящей в плане производительности из-за относительно низкой скорости передачи данных между памятью приложения и графической памятью, однако современные шины PCI Express и специальные чипы на материнской плате (например, NV3 и NV4) заметно ускорили этот процесс.

В отличие от памяти CPU, у памяти GPU есть некоторые ограничения, и доступ к ней возможен только через некоторые абстракции графического программного интерфейса. Каждая из этих абстракций может рассматриваться как поток со своим собственным набором правил доступа. Программисту доступны четыре таких потока: поток вершин, поток фрагментов, поток буфера кадров и поток текстур.

Одна из главных особенностей работы с GPU — это необходимость оптимизации кода под параллельную обработку. Это достигается использованием таких технологий, как CUDA (Compute Unified Device Architecture) от NVIDIA, OpenCL (Open Computing Language) и других. Технология CUDA представляет собой платформу для параллельных вычислений, которая позволяет программистам использовать графические ускорители NVIDIA в целях обработки больших объемов данных с высокой скоростью. Она предоставляет набор инструментов и библиотек для разработки приложений, использующих параллельные вычисления на графических процессорах. Кроме того, для эффективной работы с GPU требуется умение распараллеливать задачи и разбивать их на подзадачи, которые могут выполняться параллельно [1, 2].

При работе с GPU есть ряд преимуществ и ограничений, которые следует учитывать.

Преимущества работы с GPU:

1. Высокая скорость обработки данных. GPU может обрабатывать большое количество данных параллельно, что позволяет сократить время выполнения задачи в несколько раз по сравнению с использованием только центрального процессора.
2. Высокая энергоэффективность. GPU потребляет меньше энергии, чем CPU при выполнении тех же задач.
3. Поддержка технологий глубокого обучения. GPU используется для обучения нейронных сетей, так как он может параллельно обрабатывать множество операций с большими объемами данных.
4. Наличие специализированных библиотек и фреймворков для работы с GPU, таких как CUDA и OpenCL.



Ограничения при работе с GPU:

1. Наличие ограничений на доступную память. GPU может иметь ограничения на доступную память, что может ограничить возможности обработки больших объемов данных.
2. Невозможность использования всех типов задач. GPU может быть использован только для выполнения определенных типов задач, таких как обработка графики или параллельной обработки больших объемов данных.
3. Высокая стоимость. GPU имеет более высокую стоимость, чем CPU, что может затруднить его использование в некоторых проектах.
4. Необходимость наличия специализированных знаний для работы с GPU. Для работы с GPU необходимы специализированные знания, так как это отличается от работы с обычным CPU [3].

#### **Алгоритмы обработки данных с использованием GPU.**

Одним из способов повышения производительности обработки данных с использованием GPU является подбор оптимальных алгоритмов и структур данных. Оптимальные алгоритмы и структуры данных должны использовать максимально возможное количество ядер GPU для обработки данных, а также учитывать ограничения памяти GPU и скорость передачи данных между центральным процессором и GPU.

Алгоритмы сортировки и поиска данных — это ключевые алгоритмы в области компьютерных наук, используемые для упорядочивания и поиска данных в больших объемах информации. В зависимости от специфических требований, таких как эффективность, скорость, используемые ресурсы, существует множество различных алгоритмов, включая сортировку пузырьком, быструю сортировку, сортировку слиянием, алгоритм Шелла, алгоритмы бинарного поиска и многие другие.

Многопоточная реализация алгоритмов сортировки и поиска данных может значительно повысить их производительность и эффективность, особенно при работе с большими объемами данных. Графические ускорители могут использоваться для ускорения выполнения параллельных операций, что может быть полезным при реализации многопоточности.

Алгоритм быстрой сортировки широко известен благодаря своей стратегии "разделяй и властвуй". Он работает путем разделения массива на части относительно опорного элемента. Это позволяет переместить все элементы, которые меньше или равны опорному, влево от него, а все элементы, которые больше, - вправо. Затем происходит рекурсивное разделение каждой из частей массива, расположенных слева и справа от опорного элемента, до тех пор, пока каждая из них не будет содержать менее двух элементов.

Концепция параллельного выполнения алгоритма быстрой сортировки заключается в том, что можно одновременно сортировать подмассивы, полученные при разделении исходного массива на части вокруг опорного элемента. Для запуска параллельного процесса необходимо выполнить первое разделение массива на ЦП в последовательном режиме, чтобы установить опорный элемент и выделить подмассивы, которые будут сортироваться параллельно. Далее, технологически сортируемый массив разбивается на блоки, каждый из которых содержит часть массива, которую можно сортировать независимо от других блоков с помощью графического ускорителя. После каждого разделения степень параллелизма возрастает, т.е. увеличивается количество параллельных потоков.

Алгоритм сортировки Шелла — это усовершенствованная версия алгоритма сортировки вставками. Он основан на идее сравнения и перемещения элементов, расположенных не только рядом, но и на определенном расстоянии друг от друга в сортируемом массиве. Сам процесс сортировки включает несколько этапов, на каждом из которых несколько подписков элементов из исходного массива последовательно сортируются вставками. Например, если мы выбираем каждый  $k$ -й элемент в подписок, то получаем  $k$  подписков (шаг сортировки). Это приводит к частичной упорядоченности элементов в массиве, что увеличивает эффективность алгоритма сортировки вставками. Затем шаг сортировки уменьшается, и процедура повторяется. Окончательный результат достигается, когда шаг сортировки равен 1.

Алгоритм сортировки Шелла, адаптированный для использования на платформе CUDA, базируется на параллельном выполнении ядра функции на каждой стадии сортировки. На первой стадии величина шага выбирается так, чтобы получить максимальное количество подписков для одновременного упорядочивания в параллельных потоках, т.е. шаг =  $n / 2$  с округлением в большую сторону. Затем значение шага уменьшается вдвое, и процедура повторяется. Финальная стадия сортировки, где шаг равен 1, всегда выполняется на центральном процессоре в последовательном режиме [4].

В заключение, можно отметить, что эффективная обработка данных с использованием GPU требует подбора оптимальных алгоритмов и структур данных, а также учета особенностей конкретного задания и



доступного оборудования. Результатом правильного выбора может быть значительное повышение производительности вычислений и снижение времени обработки данных. Однако, с увеличением числа ядер и объема памяти на графических ускорителях, а также развитием инструментов для разработки приложений с поддержкой CUDA, использование данной технологии становится все более доступным и распространенным в различных сферах, включая научные и инженерные расчеты, машинное обучение, обработку изображений и видео [5, 6].

#### Экспериментальная оценка производительности.

Для проверки эффективности параллельных алгоритмов сортировки и поиска данных, реализованных с помощью программных модулей на платформе CUDA, были проведены эксперименты с прикладными программами, реализованными как последовательные на ЦП, так и многопоточные с использованием графического процессора. Измерения проводились на компьютере с процессором Intel Core i7-10700K и видеокартой NVIDIA GeForce RTX 3080.

Для сортировки использовались одномерные массивы целых чисел, сформированные случайным образом. При исследовании программ сортировки размер массива изменялся от 100000 до 1000000 элементов.

Для каждого размера массива было проведено по 10 экспериментов с разными данными. Результаты измерений времени работы алгоритмов быстрой сортировки и сортировки Шелла усреднены и представлены в таблицах 1 и 2 соответственно. Графики зависимости среднего времени сортировки от размера массива данных представлены на рисунках 1 и 2.

Таблица 1 – Среднее время работы программ быстрой сортировки

Размер массива	Последовательная реализация, мс	Параллельная реализация, мс
100 000	16	3
200 000	33	5
300 000	54	7
400 000	73	11
500 000	96	14
600 000	116	18
700 000	136	21
800 000	158	25
900 000	181	29
1 000 000	203	33

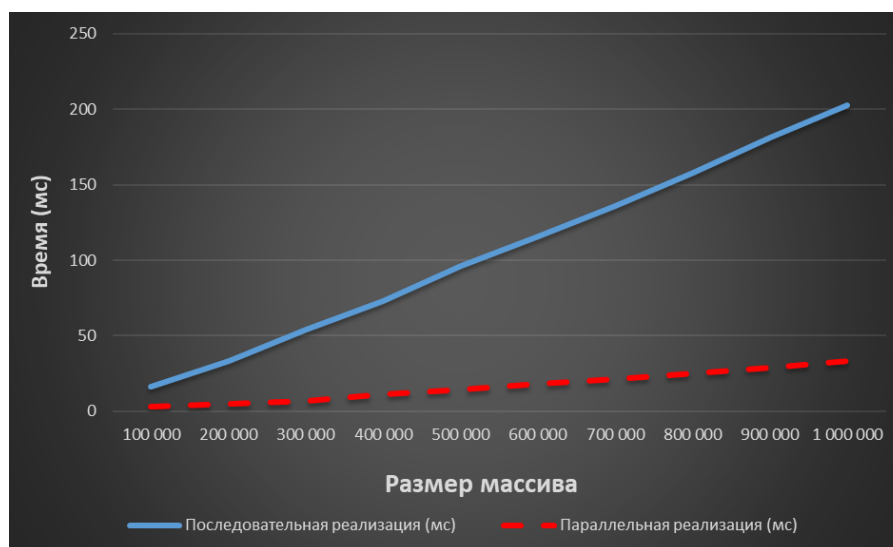


Рисунок 1. График зависимости времени работы программ быстрой сортировки от размера списка

На основании этих результатов можно сделать вывод о том, что использование графических ускорителей позволяет значительно снизить время выполнения быстрой сортировки по сравнению с последовательной реализацией на ЦП. При размере массива 1 миллион элементов время выполнения на ЦП составляет более 0.2 секунды, в то время как на GPU время выполнения составляет всего 0.033 секунды. Это позволяет существенно ускорить обработку больших объемов данных, что является важным преимуществом при работе с большими наборами данных. По таблице данных для сортировки алгоритмом Шелла можно сделать вывод, что параллельная реализация на графическом ускорителе дает значительный выигрыш во времени выполнения по сравнению с последовательной реализацией на ЦП. При увеличении размера массива данных разница во времени выполнения алгоритмов становится все более заметной. Таким образом, использование графических ускорителей для сортировки данных может значительно повысить производительность программы.

Таблица 2 – Среднее время работы программ, реализующих алгоритм Шелла

Размер массива	Последовательная реализация, мс	Параллельная реализация, мс
100 000	100	25
200 000	210	50
300 000	320	80
400 000	440	105
500 000	560	135
600 000	680	160
700 000	800	190
800 000	930	220
900 000	1050	250
1 000 000	1180	280

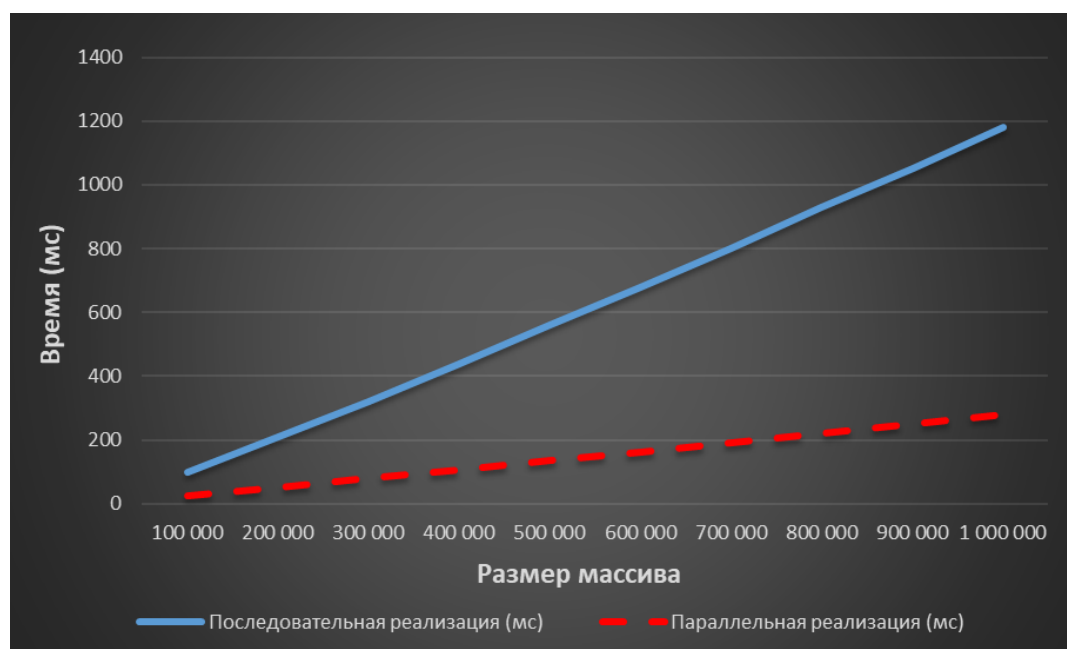


Рисунок 2. График зависимости времени работы программ алгоритма Шелла от размера списка

### Заключение.

Перспективы применения оптимизации обработки Big Data с помощью графических процессоров достаточно велики. С ростом объемов данных, которые нужно обрабатывать, становится все более важным использование эффективных методов и инструментов, которые могут обеспечить высокую скорость обработки. В этом смысле графические процессоры имеют высокий потенциал, поскольку они позволяют

параллельно обрабатывать огромные объемы данных, что значительно ускоряет процесс обработки. Кроме того, современные графические процессоры обладают высокой вычислительной мощностью, что позволяет реализовывать сложные алгоритмы обработки данных и анализа больших объемов информации. Благодаря этому применение оптимизации обработки данных с помощью GPU может быть эффективным в таких областях как наука, финансы, медицина и другие. Таким образом, оптимизация обработки больших данных с помощью графических процессоров является важной технологической задачей, которая имеет широкий спектр применения в различных областях.

#### **Список литературы**

- [1] Parallel Programming with CUDA / Cheng, J., Grossman, M., & McKercher, T. - N.-Y.: Wrox. 2018. P. 55-56  
[2] Лунин Д.В., Скворцов С.В. Организация параллельных вычислений на платформе CUDA // Вестник Рязанского государственного радиотехнического университета. 2018. № 49. С. 77-82.  
[3] Mike Houston. "General Purpose Computation on Graphics Processors" Morgan Kaufmann Publishers, 2020. - P.123  
[4] Алгоритмы и структуры данных: учеб.-метод. пособие / С. В. Актанорович, А. А. Волосевич. – Минск: БГУИР, 2018. – 112 с.  
[5] Нестеренков, С.Н. Применение технологии Big Data в игровой индустрии / С.Н. Нестеренков, М.И. Макаров, Д.В. Кишкевич // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня: сб. материалов V Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 13-14 марта 2019 года). В 2 ч. Ч. 2 / редкол.: В. А. Богуш [и др.]. - Минск: БГУИР, 2019. - С. 242-245.  
[6] Нестеренков, С. Н. Применение генетического алгоритма для решения задач многомерной оптимизации / С. Н. Нестеренков, В. Н. Наливко // Информационные технологии и системы 2019 (ИТС 2019) : материалы междунар. науч. конф., Минск, 30 окт. 2019 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2019. - С. 166-167.

## **OPTIMIZATION OF PROCESSING BIG DATA WITH GRAPHIC PROCESSING UNITS**

***S.A. Baichyk***  
*Student of BSUIR*

***S.N. Nesterenkov***  
*PhD, Associate Professor Dean of the Faculty  
of Computer Systems and Networks*

***I.S. Tarasiuk***  
*Software Engineer,  
Assistant of the Department of  
Electronic Computers*

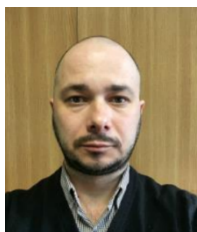
*Center for Informatization and Development of the Belarusian University of State Informatics and Radioelectronics,  
Republic of Belarus  
Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: s.bajchik@bsuir.by*

**Abstract.** This article is devoted to optimizing the processing of large volumes of data using graphics processing units (GPU). The article will provide a general overview of GPU technology, its features, and advantages in processing large data sets. The main object of the study in the article is the analysis of the performance and optimization of processing large data sets on GPU using various methods and techniques such as parallel programming, kernel and algorithm optimization, the use of specialized libraries. The aim of this work is to present and analyze different approaches to optimizing the processing of large data sets on GPU, as well as to assess the effectiveness of these approaches and their applicability for real-world conditions.

**Keywords:** Big Data, graphics processing unit, performance.

УДК 004.021:004.75

## МАТЕМАТИЧЕСКИЕ ВЫЧИСЛЕНИЯ С ИСПОЛЬЗОВАНИЕМ ГРАФИЧЕСКИХ ПРОЦЕССОРОВ ПО ТЕХНОЛОГИИ CUDA



**Е.И. Лещевич**  
аспирант БГУИР, инженер  
кафедры электронной техники  
и технологии БГУИР  
e.leshchovich@bsuir.by



**П.В. Камлач**  
заместитель декана  
факультета  
компьютерного  
проектирования, кандидат  
технических наук, доцент  
kamlachpv@bsuir.by



**В.М. Бондарик**  
декан факультета  
доуниверситетской  
подготовки и  
профессиональной ориентации,  
кандидат технических  
наук, доцент



**А.В. Чураков**  
доцент кафедры электронной  
техники и технологии,  
кан.мед.наук, врач  
анестезиолог-реаниматолог



**Г.Д. Ситник**  
доцент кафедры общей  
врачебной практики  
БелМАПО, кандидат  
медицинских наук, доцент,  
врач высшей категории по  
неврологии



**И.И. Ревинская**  
аспирант БГУИР,  
ассистент кафедры  
электронной техники и  
технологии БГУИР

### **Е.И. Лещевич**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Магистрант кафедры инженерной и компьютерной графики БГУИР. Область научного интереса – влияние инфразвука на биологические ткани.

### **П.В. Камлач**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Доцент, кандидат технических наук, заместитель декана факультета компьютерного проектирования, доцент кафедры электронной техники и технологии БГУИР. Область научного интереса – проектирование медицинских электронных систем.

### **В.М. Бондарик**

Образование: 1983-1988 Минский радиотехнический институт, специальность «Конструирование и производство радиоаппаратуры», квалификация - инженер-конструктор-технолог. Область научного интереса – проектирование медицинских электронных систем, внедрение дистанционных образовательных технологий.

### **Г.Д. Ситник**

Кандидат медицинских наук, доцент, врач высшей категории по неврологии. Область научного интереса – лечение больных с неврологическими проявлениями поясничного остеохондроза,

**И.И. Ревинская**

*Окончила Белорусский государственный университет информатики и радиоэлектроники. Аспирант кафедры электронной техники и технологии. Область научного интереса – медицинская электроника и обработка медицинских сигналов.*

**Аннотация.** Рассмотрена технология вычислений больших массивов данных на базе графических процессоров с использованием технологии CUDA.

**Ключевые слова:** CUDA, GPU, GPGPU, ALU, видеокарта, графический адаптер

**Введение.**

Устройства для превращения персональных компьютеров в маленькие суперкомпьютеры известны довольно давно. Ещё в 80-х годах прошлого века на рынке предлагались так называемые транспьютеры, которые вставлялись в распространенные тогда слоты расширения ISA. Первое время их производительность в соответствующих задачах впечатляла, но затем рост быстродействия универсальных процессоров ускорился, они усилили свои позиции в параллельных вычислениях, и смысла в транспьютерах не осталось.

Ещё несколько лет назад появились первые технологии неграфических расчётов общего назначения GPGPU (General-Purpose computation on GPUs). Ведь современные видеочипы содержат сотни математических исполнительных блоков, и эта мощь может использоваться для значительного ускорения множества вычислительно интенсивных приложений. И нынешние поколения GPU обладают достаточно гибкой архитектурой, что вместе с высокоуровневыми языками программирования и программно-аппаратными архитектурами, раскрывает эти возможности и делает их значительно более доступными.

**Актуальность.**

На создание GPCPU разработчиков побудило появление достаточно быстрых и гибких шейдерных программ, которые способны исполнять современные видеочипы. Разработчики задумали сделать так, чтобы GPU рассчитывали не только изображение в 3D приложениях, но и применялись в других параллельных расчётах. В GPGPU для этого использовались графические API: OpenGL и Direct3D, когда данные к видеочипу передавались в виде текстур, а расчётные программы загружались в виде шейдеров. Вычисления на GPU развивались и развиваются очень быстро. И в дальнейшем, два основных производителя видеочипов, Nvidia и AMD, разработали и анонсировали соответствующие платформы под названием CUDA (Compute Unified Device Architecture) и CTM (Close To Metal или AMD Stream Computing)

**Разница между CPU и GPU в параллельных расчётах.**

Рост частот универсальных процессоров упёрся в физические ограничения и высокое энергопотребление, и увеличение их производительности всё чаще происходит за счёт размещения нескольких ядер в одном чипе. Продаваемые сейчас процессоры содержат лишь до четырёх ядер (дальнейший рост не будет быстрым) и они предназначены для обычных приложений, используют MIMD — множественный поток команд и данных. Каждое ядро работает отдельно от остальных, исполняя разные инструкции для разных процессов.

Специализированные векторные возможности (SSE2 и SSE3) для четырехкомпонентных (одинарная точность вычислений с плавающей точкой) и двухкомпонентных (двойная точность) векторов появились в универсальных процессорах из-за возросших требований графических приложений, в первую очередь. Именно поэтому для определённых задач применение GPU выгоднее, ведь они изначально сделаны для них.

Например, в видеочипах Nvidia основной блок — это мультипроцессор с восемью-десятью ядрами и сотнями ALU в целом, несколькими тысячами регистров и небольшим количеством разделяемой общей памяти. Кроме того, видеокарта содержит быструю

глобальную память с доступом к ней всех мультипроцессоров, локальную память в каждом мультипроцессоре, а также специальную память для констант.

Самое главное — эти несколько ядер мультипроцессора в GPU являются SIMD (одиночный поток команд, множество потоков данных) ядрами. И эти ядра исполняют одни и те же инструкции одновременно, такой стиль программирования является обычным для графических алгоритмов и многих научных задач, но требует специфического программирования. Зато такой подход позволяет увеличить количество исполнительных блоков за счёт их упрощения.

Итак, перечислим основные различия между архитектурами CPU и GPU. Ядра CPU созданы для исполнения одного потока последовательных инструкций с максимальной производительностью, а GPU проектируются для быстрого исполнения большого числа параллельно выполняемых потоков инструкций. Универсальные процессоры оптимизированы для достижения высокой производительности единственного потока команд, обрабатывающего и целые числа и числа с плавающей точкой. При этом доступ к памяти случайный.

Разработчики CPU стараются добиться выполнения как можно большего числа инструкций параллельно, для увеличения производительности. Для этого, начиная с процессоров Intel Pentium, появилось суперскалярное выполнение, обеспечивающее выполнение двух инструкций за такт, а Pentium Pro отличился внеочередным выполнением инструкций. Но у параллельного выполнения последовательного потока инструкций есть определённые базовые ограничения и увеличением количества исполнительных блоков кратного увеличения скорости не добиться.

У видеочипов работа простая и распараллеленная изначально. Видеочип принимает на входе группу полигонов, проводит все необходимые операции, и на выходе выдаёт пиксели. Обработка полигонов и пикселей независима, их можно обрабатывать параллельно, отдельно друг от друга. Поэтому, из-за изначально параллельной организации работы в GPU используется большое количество исполнительных блоков, которые легко загрузить, в отличие от последовательного потока инструкций для CPU. Кроме того, современные GPU также могут исполнять больше одной инструкции за такт (dual issue). Так, архитектура Tesla в некоторых условиях запускает на исполнение операции MAD+MUL или MAD+SFU одновременно.

GPU отличается от CPU ещё и по принципам доступа к памяти. В GPU он связанный и легко предсказуемый — если из памяти читается текстель текстуры, то через некоторое время придёт время и для соседних текстелей. Да и при записи то же — пиксель записывается во фреймбуфер, и через несколько тактов будет записываться расположенный рядом с ним. Поэтому организация памяти отличается от той, что используется в CPU. И видеочипу, в отличие от универсальных процессоров, просто не нужна кэш-память большого размера, а для текстур требуются лишь несколько (до 128-256 в нынешних GPU) килобайт.

Да и сама по себе работа с памятью у GPU и CPU несколько отличается. Так, не все центральные процессоры имеют встроенные контроллеры памяти, а у всех GPU обычно есть по несколько контроллеров, вплоть до восьми 64-битных каналов в чипе Nvidia GT200. Кроме того, на видеокартах применяется более быстрая память, и в результате видеочипам доступна в разы большая пропускная способность памяти, что также весьма важно для параллельных расчётов, оперирующих с огромными потоками данных.

В универсальных процессорах большие количества транзисторов и площадь чипа идут на буферы команд, аппаратное предсказание ветвления и огромные объёмы начиповой кэш-памяти. Все эти аппаратные блоки нужны для ускорения исполнения немногочисленных потоков команд. Видеочипы тратят транзисторы на массивы исполнительных блоков, управляющие потоками блоков, разделяемую память небольшого

объёма и контроллеры памяти на несколько каналов. Вышеперечисленное не ускоряет выполнение отдельных потоков, оно позволяет чипу обрабатывать нескольких тысяч потоков, одновременно исполняющихся чипом и требующих высокой пропускной способности памяти.

Про отличия в кэшировании. Универсальные центральные процессоры используют кэш-память для увеличения производительности за счёт снижения задержек доступа к памяти, а GPU используют кэш или общую память для увеличения полосы пропускания. CPU снижают задержки доступа к памяти при помощи кэш-памяти большого размера, а также предсказания ветвлений кода. Эти аппаратные части занимают большую часть площади чипа и потребляют много энергии. Видеочипы обходят проблему задержек доступа к памяти при помощи одновременного исполнения тысяч потоков — в то время, когда один из потоков ожидает данных из памяти, видеочип может выполнять вычисления другого потока без ожидания и задержек.

Есть множество различий и в поддержке многопоточности. CPU исполняет 1-2 потока вычислений на одно процессорное ядро, а видеочипы могут поддерживать до 1024 потоков на каждый мультипроцессор, которых в чипе несколько штук. И если переключение с одного потока на другой для CPU стоит сотни тактов, то GPU переключает несколько потоков за один такт.

Кроме того, центральные процессоры используют SIMD (одна инструкция выполняется над многочисленными данными) блоки для векторных вычислений, а видеочипы применяют SIMT (одна инструкция и несколько потоков) для скалярной обработки потоков. SIMT не требует, чтобы разработчик преобразовывал данные в векторы, и допускает произвольные ветвления в потоках.

#### **Области применения параллельных расчётов на GPU.**

Чтобы понять, какие преимущества приносит перенос расчётов на видеочипы, приведём усреднённые цифры, полученные исследователями по всему миру. В среднем, при переносе вычислений на GPU, во многих задачах достигается ускорение в 5-30 раз, по сравнению с быстрыми универсальными процессорами. Самые большие цифры (порядка 100-кратного ускорения и даже более!) достигаются на коде, который не очень хорошо подходит для расчётов при помощи блоков SSE, но вполне удобен для GPU.

Это лишь некоторые примеры ускорений синтетического кода на GPU против SSE-векторизованного кода на CPU (по данным Nvidia):

- Флуоресцентная микроскопия: 12x;
- Молекулярная динамика (non-bonded force calc): 8-16x;
- Электростатика (прямое и многоуровневое суммирование Кулона): 40-120x и 7x.

Как видите, цифры весьма привлекательные, особенно впечатляют 100-150-кратные приросты. В следующей статье, посвящённой CUDA, мы подробно разберём некоторые из этих цифр. А сейчас перечислим основные приложения, в которых сейчас применяются вычисления на GPU: анализ и обработка изображений и сигналов, симуляция физики, вычислительная математика, вычислительная биология, финансовые расчёты, базы данных, динамика газов и жидкостей, криптография, адаптивная лучевая терапия, астрономия, обработка звука, биоинформатика, биологические симуляции, компьютерное зрение, анализ данных (data mining), цифровое кино и телевидение, электромагнитные симуляции, геоинформационные системы, военные применения, горное планирование, молекулярная динамика, магнитно-резонансная томография (MRI), нейросети, океанографические исследования, физика частиц, симуляция свёртывания молекул белка, квантовая химия, трассировка лучей, визуализация, радары, гидродинамическое моделирование (reservoir simulation), искусственный интеллект, анализ спутниковых данных, сейсмическая разведка, хирургия, ультразвук, видеоконференции.



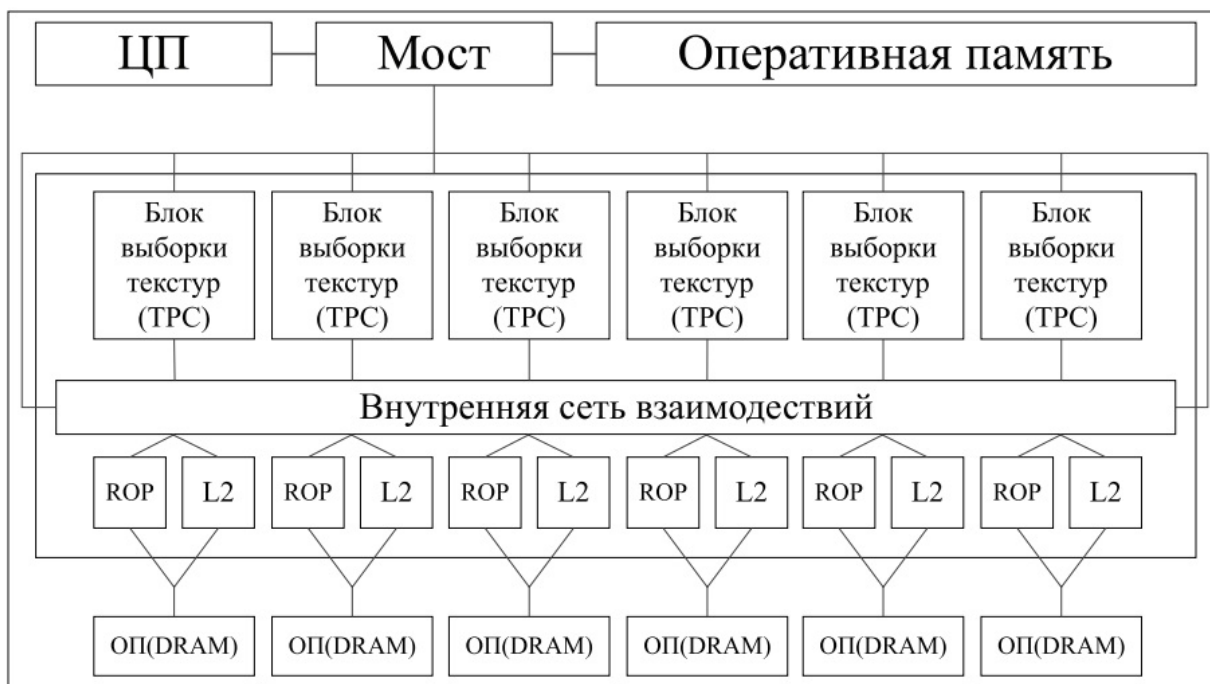


Рисунок 1. Схематическое изображение устройства графического адаптера

### Модель памяти CUDA

Модель памяти в CUDA отличается возможностью побайтной адресации, поддержкой как gather, так и scatter. Доступно довольно большое количество регистров на каждый потоковый процессор, до 1024 штук. Доступ к ним очень быстрый, хранить в них можно 32-битные целые или числа с плавающей точкой.

Каждый поток имеет доступ к следующим типам памяти (Рис. 2): глобальная память — самый большой объём памяти, доступный для всех мультипроцессоров на видеочипе, размер составляет от 256 мегабайт до 1.5 гигабайт на текущих решениях (и до 4 Гбайт на Tesla). Обладает высокой пропускной способностью, более 100 гигабайт/с для топовых решений Nvidia, но очень большими задержками в несколько сот тактов. Не кэшируется, поддерживает обобщённые инструкции load и store, и обычные указатели на память.

Локальная память — это небольшой объём памяти, к которому имеет доступ только один потоковый процессор. Она относительно медленная — такая же, как и глобальная.

Разделяемая память — это 16-килобайтный (в видеочипах нынешней архитектуры) блок памяти с общим доступом для всех потоковых процессоров в мультипроцессоре. Эта память весьма быстрая, такая же, как регистры. Она обеспечивает взаимодействие потоков, управляется разработчиком напрямую и имеет низкие задержки. Преимущества разделяемой памяти: использование в виде управляемого программистом кэша первого уровня, снижение задержек при доступе исполнительных блоков (ALU) к данным, сокращение количества обращений к глобальной памяти.

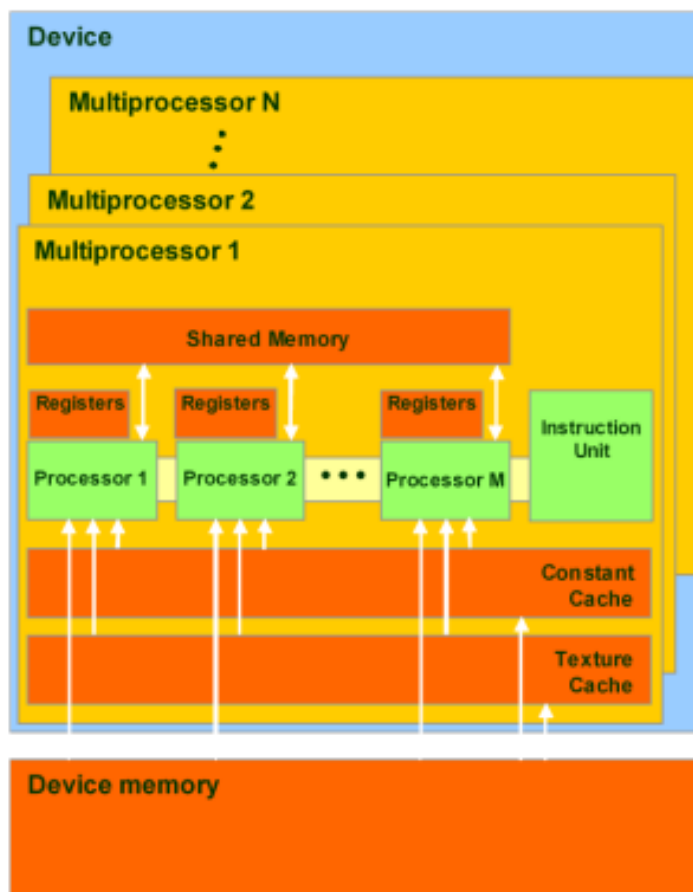


Рисунок 2. Модель доступа к различным типам памяти

Память констант — область памяти объемом 64 килобайта (то же — для нынешних GPU), доступная только для чтения всеми мультипроцессорами. Она кэшируется по 8 килобайт на каждый мультипроцессор. Довольно медленная — задержка в несколько сот тактов при отсутствии нужных данных в кэше.

Текстурная память — блок памяти, доступный для чтения всеми мультипроцессорами. Выборка данных осуществляется при помощи текстурных блоков видеочипа, поэтому предоставляются возможности линейной интерполяции данных без дополнительных затрат. Кэшируется по 8 килобайт на каждый мультипроцессор. Медленная, как глобальная — сотни тактов задержки при отсутствии данных в кэше.

Естественно, что глобальная, локальная, текстурная и память констант — это физически одна и та же память, известная как локальная видеопамять видеокарты. Их отличия в различных алгоритмах кэширования и моделях доступа. Центральный процессор может обновлять и запрашивать только внешнюю память: глобальную, константную и текстурную.

Из написанного выше понятно, что CUDA предполагает специальный подход к разработке, не совсем такой, как принят в программах для CPU. Нужно помнить о разных типах памяти, о том, что локальная и глобальная память не кэшируется и задержки при доступе к ней гораздо выше, чем у регистровой памяти, так как она физически находится в отдельных микросхемах.

Типичный, но не обязательный шаблон решения задач:

- задача разбивается на подзадачи;
- входные данные делятся на блоки, которые вмещаются в разделяемую память;

- каждый блок обрабатывается блоком потоков;
- подблок подгружается в разделяемую память из глобальной;
- над данными в разделяемой памяти проводятся соответствующие вычисления;
- результаты копируются из разделяемой памяти обратно в глобальную.

### Среда программирования

В состав CUDA входят runtime библиотеки:

- общая часть, предоставляющая встроенные векторные типы и подмножества вызовов RTL, поддерживаемые на CPU и GPU;
- CPU-компонента, для управления одним или несколькими GPU;
- GPU-компонента, предоставляющая специфические функции для GPU.

Основной процесс приложения CUDA работает на универсальном процессоре (host), он запускает несколько копий процессов kernel на видеокарте. Код для CPU делает следующее: инициализирует GPU, распределяет память на видеокарте и системе, копирует константы в память видеокарты, запускает несколько копий процессов kernel на видеокарте, копирует полученный результат из видеопамяти, освобождает память и завершает работу.

В качестве примера для понимания приведем CPU код для сложения векторов, представленный в CUDA:

```
//Размер вектора в элементах
const int N = 1048576;
//размер вектора в байтах
const int dataSize = N * sizeof(float);

//Выделение памяти CPU
float *h_A = (float *)malloc(dataSize);
float *h_B = (float *)malloc(dataSize);
float *h_C = (float *)malloc(dataSize);

//Выделение памяти GPU
float *d_A, *d_B, *d_C;
cudaMalloc((void **)&d_A, dataSize);
cudaMalloc((void **)&d_B, dataSize);
cudaMalloc((void **)&d_C, dataSize);

//Инициализировать h_A[], h_B[]...

//Скопировать входные данные в GPU для обработки
cudaMemcpy(d_A, h_A, dataSize, cudaMemcpyHostToDevice);
cudaMemcpy(d_B, h_B, dataSize, cudaMemcpyHostToDevice);

//Запустить ядро из N / 256 блоков по 256 потоков
//Предполагая, что N кратно 256
vectorAdd<<<N / 256, 256>>>(d_C, d_A, d_B);

//Считать результаты GPU
cudaMemcpy(h_C, d_C, dataSize, cudaMemcpyDeviceToHost);
```

Рисунок 3. Пример кода сложения векторов

Функции, исполняемые видеочипом, имеют следующие ограничения: отсутствует рекурсия, нет статических переменных внутри функций и переменного числа аргументов.

Поддерживаются два вида управления памятью: линейная память с доступом по 32-битным указателям, и CUDA-массивы с доступом только через функции текстурной выборки.

Программы на CUDA могут взаимодействовать с графическими API: для рендеринга данных, сгенерированных в программе, для считывания результатов рендеринга и их обработки средствами CUDA (например, при реализации фильтров постобработки). Для этого ресурсы графических API могут быть отображены (с получением адреса ресурса) в пространство глобальной памяти CUDA. Поддерживаются следующие типы ресурсов графических API: Buffer Objects (PBO / VBO) в OpenGL, вершинные буферы и текстуры (2D, 3D и кубические карты) Direct3D9.

Функции, исполняемые видеочипом, имеют следующие ограничения: отсутствует рекурсия, нет статических переменных внутри функций и переменного числа аргументов. Поддерживаются два вида управления памятью: линейная память с доступом по 32-битным указателям, и CUDA-массивы с доступом только через функции текстурной выборки.

Программы на CUDA могут взаимодействовать с графическими API: для рендеринга данных, сгенерированных в программе, для считывания результатов рендеринга и их обработки средствами CUDA (например, при реализации фильтров постобработки). Для этого ресурсы графических API могут быть отображены (с получением адреса ресурса) в пространство глобальной памяти CUDA. Поддерживаются следующие типы ресурсов графических API: Buffer Objects (PBO / VBO) в OpenGL, вершинные буферы и текстуры (2D, 3D и кубические карты) Direct3D9.

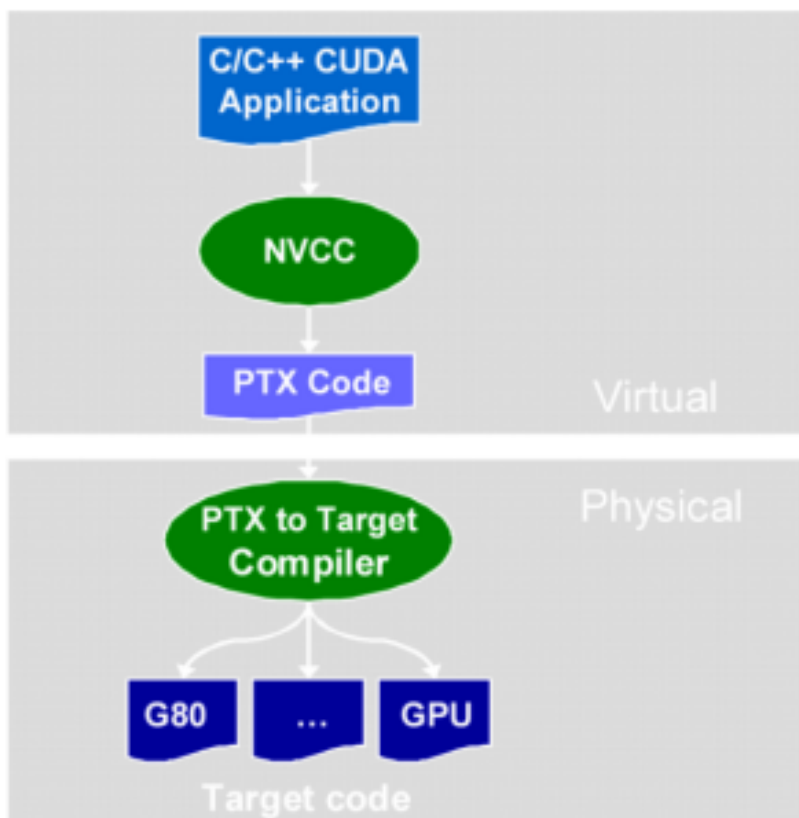


Рисунок 4. Стадии компиляции CUDA-приложения

Файлы исходного кода на CUDA C компилируются при помощи программы NVCC, которая является оболочкой над другими инструментами, и вызывает их: cudacc, g++, cl и

др. NVCC генерирует: код для центрального процессора, который компилируется вместе с остальными частями приложения, написанными на чистом Си, и объектный код PTX для видеочипа. Исполнимые файлы с кодом на CUDA в обязательном порядке требуют наличия библиотек CUDA runtime library (cudart) и CUDA core library (cuda).

#### **Заключение.**

Представленная компанией Nvidia программно-аппаратная архитектура для расчётов на видеочипах CUDA хорошо подходит для решения широкого круга задач с высоким параллелизмом. CUDA работает на большом количестве видеочипов Nvidia, и улучшает модель программирования GPU, значительно упрощая её и добавляя большое количество возможностей, таких как разделяемая память, возможность синхронизации потоков, вычисления с двойной точностью и целочисленные операции.

#### **Список литературы**

- [1] Параллельные вычисления на GPU. Архитектура и программная модель CUDA: Учебное пособие. А. В. Боресков и др. Предисл.: В. А. Садовничий. Издательство Московского университета, 2012
- [2] Технология CUDA в примерах. Введение в программирование графических процессоров Джейсон Сандерс, Эдвард Кэндрот ДМК Пресс, 2011 г.
- [3] Боресков А. В., Харламов А. А., Основы работы с технологией CUDA, Москва: ДМК Пресс, 2010.
- [4] Краткий курс обработки изображений. -URL: <https://hub.exponenta.ru/post/kratkiy-kurs-teorii-obrabotkiizobrazheniy734>

## **MATHEMATICAL CALCULATIONS USING GRAPHIC PROCESSORS USING CUDA TECHNOLOGY**

### ***E.I. Leshchevich***

*Postgraduate student of the BSUIR, assistant of the Department of Electronic Engineering and Technology of BSUIR*

### ***P.V. Kamlach***

*Deputy Dean of of Computer-Aided Design, PhD, Associate Professor at the Department of Electronic Technique and Technology*

### ***V. M. Bandarik***

*Dean of the Faculty of Pre-University Preparation and Occupational Guidance, PhD, Associate Professor*

### ***A.V. Churakov***

*PhD, Associate Professor at the Department of Electronic Technique and Technology*

### ***G.D. Sitnik***

*PhD, Associate Professor at the Department of Belarusian Medical Academy of Postgraduate Education, doctor of the highest category in neurology*

### ***I.I. Revinskaya***

*Postgraduate at the Department of Electronic Technique and Technology*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus  
Email: e.leshchevich@bsuir.by, kamlachpv@bsuir.by*

**Abstract.** The technology of computing large data arrays based on graphic processors using CUDA technology is considered.

**Keywords:** CUDA, GPU, GPGPU, ALU, graphics card, graphics adapter

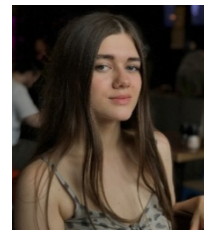
УДК 339.138:004.7

## ИСПОЛЬЗОВАНИЕ ВЕБ-АНАЛИТИКИ ДЛЯ РЕШЕНИЯ БИЗНЕС-ЗАДАЧ КОРПОРАТИВНОГО САЙТА



**Д.А. Фролова**

Преподаватель кафедры экономики БГУИР,  
магистр экономических наук  
frolova.profstud@gmail.com



**Ю.А. Янукович**

Студентка инженерно-  
экономического факультета БГУИР  
pinkrest19@gmail.com

### **Д.А. Фролова**

Окончила БГУИР. Область научных интересов связана с механизмами влияния маркетинговых возможностей на формирование и развитие конкурентного бизнеса, нейромаркетинг.

### **Ю.А. Янукович**

Обучается в БГУИР. Область научных интересов связана с современными инструментами интернет-маркетинга, стратегиями развития электронного бизнеса.

**Аннотация.** В данной статье рассмотрена веб-аналитика как эффективный способ анализа и оценки эффективности деятельности компании в сети Интернет. Проанализированы последние тенденции использования веб-аналитики в мире, а также влияние использования аналитических средств на успешность функционирования предприятия на электронном рынке. Приведены основные этапы веб-анализа корпоративных сайтов и способы решения ключевых бизнес-задач.

**Ключевые слова:** веб-аналитика, решение бизнес-задач, оптимизация сайта.

### **Введение.**

Digital-технологии в бизнес-процессах сегодня являются естественным решением для эффективной работы компании. Веб-аналитика — это сбор, синтез и анализ данных веб-сайта с целью улучшения его функционирования и качества взаимодействия с пользователями. Компании используют платформы веб-аналитики для измерения и оценки производительности сайта, а также для просмотра ключевых показателей эффективности. Также платформы веб-аналитики оценивают активность и поведение пользователей на веб-сайте, исследуя количество посещений, время, проведенное на сайте, просматриваемые страницы.

Учитывая тенденцию увеличения частоты использования смартфонов при просмотре различных сайтов, повышается качество получаемой, детализированной информации, которая используется в аналитических системах. Соответственно, это формирует устойчивый рост и развитие рынка веб-аналитики. За последнее время годовой темп роста данного рынка составил 19,72%, так к 2030 году прогнозируется увеличение размера рынка и использования систем веб-аналитики более чем в 3,5 раза [1]:

### **Способы решения бизнес-задач системами веб-аналитики.**

Использование систем веб-аналитики актуально для всех компаний, присутствующих в цифровом пространстве. При помощи современных интернет-инструментов компании могут выполнять расширенный анализ различных данных, даже без привлечения программной специалистов по обработке и анализу данных. Предприятия обретают новое конкурентное преимущество и возможность выявить тенденции для дальнейшего развития бизнеса. Системы веб-аналитики помогают решить следующие задачи компаний [2]:





Рисунок 1. Динамика развития мирового рынка Web-аналитики

1) Исследование целевой аудитории. Инструменты веб-аналитики раскрывают ключевые сведения о посетителях вашего сайта, в том числе о среднем времени, проведенном ими на странице, о том, являются ли они новыми или вернувшимися пользователями, а также о том, какой контент привлекает больше всего трафика. С помощью этой информации можно узнать интересы потенциальных клиентов и составить их портрет.

2) Бюджетирование. В комплексной стратегии маркетинга важно рассматривать все каналы дохода и оценивать их эффективность. С данной задачей помогает разобраться аналитика, просматривая все источники трафика, дохода, учитывая наиболее активные сегменты пользователей и будущие прогнозы, составляется наиболее рациональный план распределения бюджета.

3) Оптимизация веб-сайта. Детальный просмотр пути клиента, который проходит по страницам входа, выхода, находит ключевые страницы конверсий, помогает компаниям найти возможные технические ошибки, недочеты в размещаемом контенте и, наоборот, точки роста. Одними из самых используемых систем веб-аналитики в настоящее время являются Яндекс.Метрика и Google Analytics. При помощи них проводят общий анализ работы сайта, оценивают поведение аудитории и рассчитывают эффективность проведенных рекламных кампаний. Установка аналитических систем позволят ответить на один из самых важных вопросов: "Какие действия происходят после перехода на сайт посетителя?".

#### **Ключевые этапы веб-аналитики.**

Для того, чтобы инструменты веб-аналитики работали эффективно и приносили значимые результаты, важно соблюдать такие ключевые этапы [3]:

1) Выбор цели. Необходимо выбрать цели, которые должны быть достигнуты. Это может быть увеличение продаж, повышение узнаваемости компании. Бизнес-цели могут оцениваться как количественно, так и качественно.

2) Поиск данных. Второй шаг в веб-аналитике — это поиск и накопление аналитических данных. Данные могут быть собраны с корпоративных веб-сайтов или через системы веб-аналитики (Яндекс.Метрика и Google Analytics).

3) Обработка данных. На следующем этапе происходит обработка, распределение всех данных по сегментам (например, по странам, этапам воронки продаж).

4) Выбор KPI. В веб-аналитике KPI — это показатели для отслеживания и оценки поведения потенциальных покупателей на сайте. Например: количество уникальных пользователей, каналы трафика, конверсии.

5) Разработка стратегии. Этот этап включает в себя внедрение идей для формулирования стратегий, которые соответствуют целям организации. Например, поисковые запросы, выполняемые на сайте, могут помочь организации разработать контент-стратегию на основе того, что пользователи ищут на ее веб-сайте.



б) Испытание стратегий. Компании должны пробовать работать с различными стратегиями, чтобы найти ту, которая самая эффективная. К примеру, А/В-тестирование может помочь узнать, какой контент может больше привлекать аудиторию. В процессе создаются два и более варианта контента, а затем они показываются различным сегментам целевой аудитории, чтобы понять, какая из стратегий подходит лучше.

#### **Заключение.**

Следует отметить, что веб-аналитика может стать эффективным маркетинговым инструментом для любой сферы бизнеса. Аналитические системы помогут улучшить качество сайта компании, увеличить долю клиентов среди потребителей, оценить эффективность рекламных кампаний, повысить качество обслуживания посредством формирования успешной стратегии развития в сети Интернет. Организации оценивают свой текущий уровень развития в Интернете на основании получаемых данных, экспериментируют, находят необходимое направление для роста и, как итог, повышают свою конкурентоспособность и прибыльность.

#### **Список литературы**

[1] Market Research Future – Web Analytics Market Report - Global Forecast to 2030 [Электронный ресурс]. – URL. <https://www.marketresearchfuture.com/reports/web-analyticst> (дата обращения 16.03.2023).

[2] Amplitude – What is Web Analytics? [Электронный ресурс]. – URL. <https://amplitude.com/blog/web-analytics> (дата обращения 16.03.2023).

[3] TechTarget Contributor – Web Analytics [Электронный ресурс]. – URL. <https://www.techtarget.com/searchbusinessanalytics/definition/Web-analytics> (дата обращения 17.03.2023).

## **USING WEB ANALYTICS TO SOLVE BUSINESS PROBLEMS OF A CORPORATE WEBSITE**

***D.A. Frolova,***

*Lecturer*

*BSUIR Economics, Master of Economic Sciences*

***Y.A. Yanukovich***

*Student of BSUIR*

*Belarusian State University of Informatics and Radioelectronics*

*Minsk, Republic of Belarus*

*E-mail: frolova.profstud@gmail.com, pinkerest19@gmail.com*

**Abstract.** This article discusses web analytics as an effective way to analyze and evaluate the effectiveness of a company's activities on the Internet. The latest trends in the use of web analytics in the world are analyzed, as well as the impact of the use of analytical tools on the success of the enterprise in the electronic market. The main stages of web analysis of corporate websites and ways of solving key business problems are given.

**Keywords:** web analytics, solving business problems, website optimization.

УДК 004.021:004.75

## МОДЕЛИРОВАНИЕ МЕТОДА СЖАТИЯ И ОБРАТНОГО ВОССТАНОВЛЕНИЯ ДВОИЧНОГО КОДА С ИСПОЛЬЗОВАНИЕМ ПОЛИНОМИАЛЬНЫХ КОРРЕЛИРОВАННЫХ ФУНКЦИЙ



**В.П. Домеников**

Директор ООО «Красное Солнце», PhD в области информационных технологий (радиосвязь), аспирант БГУИР  
[domenicoff@gmail.com](mailto:domenicoff@gmail.com)



**А.Г. Сапëров**

Академик МНОО(МАИТ), доктор наук в области информационных технологий, профессор, кандидат технических наук  
[carel@tut.com](mailto:carel@tut.com)



**А.С. Строгова**

БГУ, Заместитель начальника ГУН-начальник отдела организации и сопровождения инновационной деятельности, кандидат технических наук, доцент  
[strogova@gmail.com](mailto:strogova@gmail.com)

### **В.П. Домеников**

Окончил Могилевский государственный университет имени А.А. Кулешова. Область научных интересов связана с разработкой методов и алгоритмов защиты и сжатия информации.

### **А.Г. Сапëров**

Окончил Могилевский машиностроительный институт. Область научных интересов связана с исследованием проблем пропускной способности каналов и кодирования информации.

### **А.С. Строгова**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с нанотехнологиями и наноматериалами, элементы памяти.

**Аннотация.** Произведен анализ существующих методов и алгоритмов сжатия последовательного двоичного кода, обеспечивающих однозначное и без потерь его восстановление.

Показано, что при аппроксимации закодированного битового потока коррелированными полиномиальными функциями наступает эффект сжатия. Приведен алгоритм и разработана схема восстановления исходного кода с помощью интерполятора, используя мажоритарный принцип.

**Ключевые слова:** аппроксиматор, интерполятор, корреляция, полином, мажоритарный принцип, компаратор.

### **Введение.**

Существует достаточно много алгоритмов и схемотехнических решений для реализации сжатия и восстановления без потерь информации [6]. Все они классифицируются по определенным признакам и назначению, в зависимости от предъявляемых требований к характеристикам компрессии. В настоящей статье рассмотрен алгоритм и схемная модель, позволяющая кодировать исходный последовательный двоичный код специальными функциями. Это позволяет в дальнейшем, при организации сжатия битового потока, применить алгоритм безошибочного восстановления.

### **Актуальность.**

Оценка сжатия и обратного восстановления данных является одним из наиболее перспективных и актуальных направлений исследования для современных систем хранения информации. Это обусловлено тем, что, например, при передаче данных затраты времени вступают в конфликт с оперативностью использования последних. Можно привести ряд примеров, когда принятая информация становится уже не актуальной в виду потери времени.

### Сущность предлагаемого метода.

В основу метода сжатия положен принцип аппроксимации цифрового потока данных оригинала [1-5]. Исходный последовательный двоичный код (ПДК) кодируется коррелированными полиномиальными функциями вида:

$$Y(j \cdot T) = [j(j+1)/2]^2, \quad (1)$$

где  $Y(j \cdot T)$  – значение цифрового отсчета функции;

$j$  – порядковый номер отсчета;

$T$  – интервал времени, с которым следует функция.

ПДК кодируется в соответствии с принципом суперпозиции по приведенной ниже схеме [2]

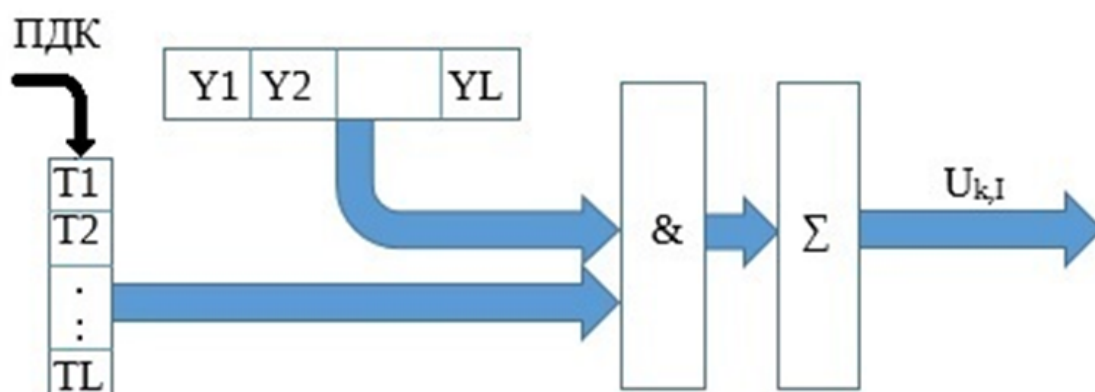


Рисунок 1. Структурная схема для кодирования ПДК.

Условные обозначения на схеме:

$T1 - TL$  – триггера регистра сдвига;

$L$  – длина интервала корреляции;

$Y1 = Y(1 \cdot T), Y2 = Y(2 \cdot T), YL = Y(L \cdot T) \dots$  - двоичные коды цифровых отсчетов функции;

$\&$  - перемножители;

$\Sigma$  - пирамидальный сумматор;

$U_{k,l}$  – результирующее значение суперпозиции

$$U_{k,l} = \sum Y(j \cdot T) \cdot X_{i+j+1}, \quad (2)$$

где  $X_i$  – текущее значение разряда ПДК.

В результате работы схемы на рисунке 1 на выходе пирамидального сумматора  $\Sigma$  будет реализован цифровой поток данных в соответствии с формулой (2). Для организации процесса сжатия ПДК используется алгоритм, состоящий из 5 процедур и выполняемых в порядке их перечисления.

Процедура 1 – детерминированная аппроксимация цифрового потока  $U_{k,l}$ .

Процедура 2 – извлечение из таблицы декодирования (ТД) двоичного кода, соответствующего заданному значению  $U_{k,l}$ .

Процедура 3 – накопление результатов декодирования по ТД в блоке счетчиков в соответствии с мажоритарным принципом.

Процедура 4 – сдвиг влево (перезапись) содержимого каждого счетчика.

**Процедура 5** – определение истинного значения ПДК с помощью компаратора:

$X_i = \langle 1 \rangle$ , если значение крайнего левого счетчика  $> L/2$ ;

$X_i = \langle 0 \rangle$ , если значение крайнего левого счетчика  $\leq L/2$ .

ТД имеет вид (для  $L=3$ ) [2]:

Таблица 1. Таблица декодирования цифровых отсчетов при  $L=3$ .

№п/п	$X_i$	$X_{i-1}$	$X_{i-2}$	$U_{k,i}$
1	0	0	0	0
2	1	0	0	1
3	0	1	0	9
4	1	1	0	10
5	0	0	1	36
6	1	0	1	37
7	0	1	1	45
8	1	1	1	46

Реализация вышеперечисленных процедур алгоритма осуществляется представленной на рисунке 2 структурной схемой:

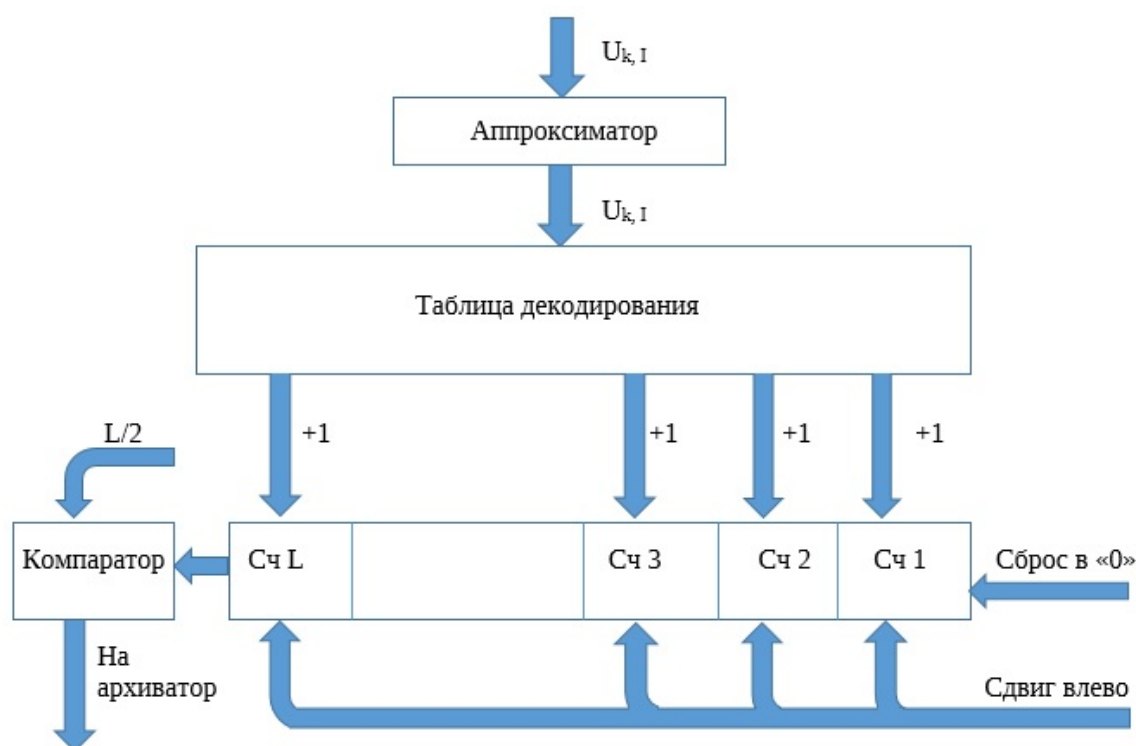


Рисунок 2. Схема компрессора ПДК

Условное обозначение на схеме: Сч1—СчL – блок счетчиков из L счетчиков.

Обратное восстановление в оригинальный поток данных возможно, если будут восстановлены аппроксимированные детерминированные значения закодированного цифрового потока.

Поскольку речь идет о коррелированных значениях ПДК, то абсолютные значения интерполируемых значений  $U_{k,i}$  должны не превышать максимального динамического диапазона функции  $Y(j \cdot T)$  (1).

Вследствие этого необходимым и достаточным условием восстановления скомпрессированного ПДК будет выполнение последовательно трех процедур:

- 1- реализация работы схемы на рисунке 1;
- 2- интерполяция (восстановление) отсчетов в сжатом потоке данных;
- 3- работа схемы на рисунке 3.

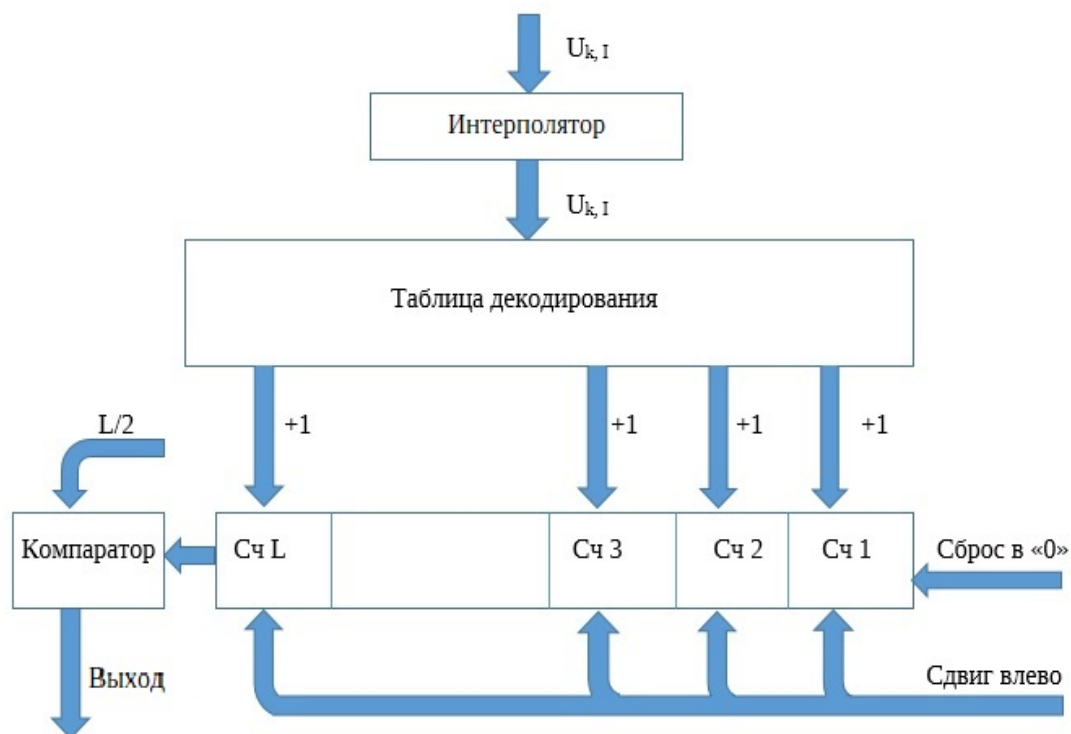


Рисунок 3. Структурная схема декомпрессора

### Заключение.

Использование предложенной методики сжатия двоичного кода позволяет перейти к компрессии файлов, имеющих максимально возможную энтропию.

В отличие от существующих алгоритмов и методов, мажоритарный принцип дает возможность применять для сжатия и восстановления без потерь достаточно простые математические методы аппроксимации и интерполяции.

### Список литературы

- [1] Трубицын, Л.М. Передача аналогового сигнала перекрывающимися импульсами / Л.М. Трубицын // Изв. высш.учеб. заведений. Приборостроение. – 1986. – №. 6.
- [2] Устройство для сжатия данных с их последующим восстановлением / А.Г. Саперов, В.П. Домеников, Н.Н. Уласюк // Патент РБ №23194 от 01.09.2020г.
- [3] Saperov, A. G. The Method for repeated compressing of data and the multiprocessor convertor / A.G. Saperov // WIPO/PCT. Application for utility patent. International Publication Number WO 2016/185254 A1. – International Publication Date November. 24. – 2016.
- [4] Сапёров А.Г., Домеников В.П. Преобразователь последовательного двоичного кода в десятичный. – Патент РБ № 23531 от 30.07.2021г.
- [5] Способ защиты информации и пирамидальный криптографический процессор / А.Г. Сапёров, В.П. Домеников // Патентная заявка РБ, а 201900065 от 07.03.2019 г. Официальный бюллетень № 3 от 30.06.2020.
- [6] Лезин, Ю.С. Введение в теорию и технику радиотехнических систем / под ред. Ю.С. Лезин – М.: Радио и связь, 1986 – 280 с.

## **MODELING THE COMPRESSION AND BACKRECOVERY METHOD OF BINARY CODE USING POLYNOMIAL CORRELATED FUNCTIONS**

***V.P. Domenikof***

*Director of LLC «Krasnoe solnce»,  
PhD of Information Technology  
(radio communication), postgraduate  
student of BSUIR.*

***A.G. Saperov***

*Academician of INTERNATIONAL  
SCIENTIFIC PUBLIC UNION  
"IAIT", Doctor of Science in  
Information Technology, Professor,  
Candidate of Technical Sciences.*

***A.S. Strogova***

*Deputy Head of the General  
Directorate of Science-Head of the  
Department of Organization and  
support of innovative Activities, PhD  
of Technical Sciences, Associate  
Professor.*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: alexvikt.minsk@gmail.com*

**Abstract.** An analysis of the existing methods and algorithms for compressing a serial binary code, providing its unambiguous and lossless recovery, is carried out.

It is shown that when approximating the encoded bit stream by correlated polynomial functions, the effect of compression occurs. An algorithm is presented and a scheme for restoring the source code using an interpolator using the majority principle is developed.

**Keywords:** approximator, interpolator, correlation, polynomial, majority principle, comparator.

УДК 004.62:316.472.4

## BIG DATA В ВЕБ ПРИЛОЖЕНИИ ПО КОНТРОЛЮ ВЫПОЛНЕННОЙ РАБОТЫ В КОМАНДЕ



**П.А. Такунов**

Студент 4 курса факультета компьютерных систем и сетей специальности ИиТП БГУИР  
[pavel.takunov@gmail.com](mailto:pavel.takunov@gmail.com)



**С.Н. Нестеренков**

Декан факультета компьютерных систем и сетей БГУИР, кандидат технических наук, доцент  
[s.nesterenkov@bsuir.by](mailto:s.nesterenkov@bsuir.by)



**А.Н. Марков**

Старший преподаватель факультета компьютерных систем и сетей кафедры ПОИТ БГУИР  
[a.n.markov@bsuir.by](mailto:a.n.markov@bsuir.by)

### **П.А. Такунов**

Студент 4 курса специальности «Информатика и Технологии Программирования» Белорусского государственного университета информатики и радиоэлектроники.

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцента кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации

### **А.Н. Марков**

Старший преподаватель кафедры ПОИТ, заместитель начальника ЦИИР.

**Аннотация.** Web-приложение по контролю выполненной работы в команде - это высокоэффективное средство управления проектами, которое позволяет следить за выполняемыми задачами, оценивать эффективность использования времени и ресурсов, а также координировать работу всей команды.

Оно позволяет создавать проекты, присваивать им задачи и делить их между участниками команды. Пользователи могут мониторить статус задач, отслеживать прогресс и сроки выполнения работ, а также обмениваться комментариями и файлами.

Приложение оснащено функциями анализа продуктивности команды, а также возможностью вносить коррективы в проект, если сроки реализации сбиваются с графика.

**Ключевые слова:** веб-приложение, контроль выполнения работ, командная работа, управление проектами, анализ продуктивности, сроки выполнения работ, анализ данных, визуализация информации, Big Data.

### **Введение.**

Одним из самых больших недостатков удалённой работы является доверие: менеджеры просто не доверяют своим людям работать без присмотра. Они привыкли управлять, подсчитывая людей на местах, а не результаты.

Это не управление, надзора за сотрудниками.

Более того, присутствие человека на рабочем месте ничего не говорит менеджеру о том, действительно ли этот человек работает. Когда клиенты спрашивают: «Как я узнаю, работают ли они?» Я спрашиваю: «Откуда вы знаете, что они сейчас работают?» Эксперты по менеджменту превозносят необходимость управления по результатам уже более четырех десятилетий. Микроуправление не работает, равно как и «управление путем обхода» в этом глобальном, мобильном мире. Если люди будут вынуждены работать дома в течение длительного периода времени, а, похоже, так и будет, менеджерам придётся понять, что важны результаты.



### Актуальность.

За последние несколько лет основной движущей силой программ работы на дому было привлечение и удержание талантов, но во время последней рецессии речь шла в основном об экономии денег. Руководители организаций, отчаянно пытавшиеся сократить расходы, обнаружили, что могут добиться большего, используя меньше недвижимости.

С тех пор исследования занятости показали, насколько неэффективно использовались офисные площади. Сотрудники по всему миру не находятся за рабочим столом от 50 до 60 % времени. Это огромная трата денег.

Согласно данным компании Global Workplace Analytics, количество работающих удаленно в США выросло на 159% с 2005 года и на 44% с 2015 года. Их последний отчет «State of Remote Work 2021» показал, что на начало 2021 года 42% американских работников работали удаленно (в сравнении с 30% в 2019 году) [1].

По данным исследования рекрутинговой компании HeadHunter, среди российских компаний, особенно крупных, поддерживающих удаленную работу, ее доля выросла в три раза с конца 2019 года. Так, 72% крупнейших компаний были готовы работать удаленно в апреле 2020 года, а в конце года этот показатель вырос до 91% [2].

Интернет-журнал «РБК» пишет о том, что в России до пандемии удаленно работало около 2% работающих, а к середине апреля 2020 года — уже около 30%.

В целом, в мире наблюдается стремительный рост количества работающих удаленно на фоне пандемии COVID-19. В некоторых странах доля таких работников достигла 50-60%.

Даже после спада активности вируса в мире люди не спешат возвращаться в офисы. Согласно проведенному опросу получаем следующие результаты, представленные на рисунке 1.



Рисунок 1. Опрос «Почему люди хотят работать удаленно» [3]

Из которых видно, что эффективность рабочих увеличилась и стрессовых ситуаций стало меньше. А работодатели экономят на помещённых.

Основными недостатками при переходе на удалённую работу являются:

1. Зависимость от интернета: веб-приложения требуют подключения к Интернету, что может быть неудобно в зоне низкой скорости интернета или в местах с плохой связью.

2. Безопасность: хранение данных в облаке может быть небезопасным и приложения могут стать жертвами хакерских атак или утечек данных.

3. Сложность в использовании: некоторые веб-приложения могут иметь сложный интерфейс, что может быть сложно для непрофессиональных пользователей.

4. Высокая стоимость: использование веб-приложений может быть довольно дорогим, особенно для малых компаний.

Также есть люди, которые не хотят переходить на удалённую работу. Среди таких людей тоже был проведён опрос и результаты можно увидеть на рисунке 2.

## ПРОБЛЕМЫ ПРИ ОРГАНИЗАЦИИ УДАЛЕННОЙ РАБОТЫ

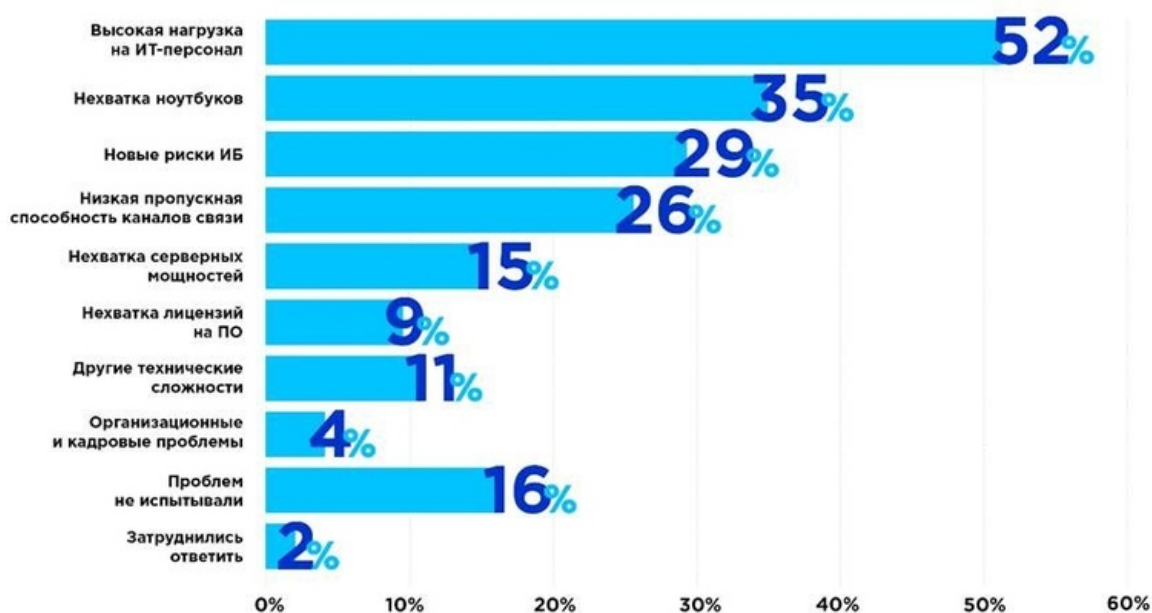


Рисунок 2. Опрос «Почему люди не хотят работать удаленно» [3]

### Основная часть.

Определим какую ценность и влияние имеют большие данные для команды.

Помимо всего прочего, большие данные могут помочь инициаторам команды получить ответы на наиболее важные вопросы, а также получить необходимую информацию по таким темам, как:

- Нужно ли нанимать новых сотрудников?
- Нужно ли уволить сотрудников?
- Какова эффективность сотрудников?
- Нужно ли увеличить или уменьшить зарплату?
- Нужно ли сокращение сотрудников?
- Сколько времени требуется на выполнение работы?

Веб приложение по контролю выполненной работы в команде - это эффективный инструмент управления проектами, который позволяет улучшить совместную работу и сократить время выполнения задач.

Однако, современные технологии, такие как анализ больших данных, машинное обучение и визуализация информации могут внести еще большую ценность в этот процесс.

Большие данные могут быть полезны в анализе веб-приложений по контролю выполненной работы в команде, так как с помощью алгоритмов машинного обучения и анализа данных можно более точно и полно анализировать продуктивность команды и эффективность использования времени и ресурсов.

Например, большие данные могут использоваться для определения наиболее эффективных способов закрепления участников команды за выполненными задачами, оптимизации длительности работ и координации членов команды.

Также, анализ больших данных может помочь определить тенденции в области выполнения задач и проектов, что может быть полезно при разработке новых проектов и в будущей работе команды.

В целом, использование данных и аналитики в сфере управления проектами может увеличить эффективность работы команды и успех кампаний.

Продвинутое веб-приложение, основанное на этом подходе, позволяет выявлять слабые места в проекте, улучшать его продуктивность и прогнозировать результаты. Оно также позволяет командам работать вместе более эффективно и сокращать время, затраченное на задачи.

Для оценки зачастую используются следующие методы:

— Методология Agile - позволяет команде быстро адаптироваться к изменениям в проекте и эффективно работать.

— Методология Six Sigma - используется для оценки качества работы команды, основываясь на статистических данных. Формула для расчета процесса Six Sigma: (Число стандартных отклонений от среднего  $\times$  Уровень уверенности) + Среднее значение (стандартное отклонение  $\times$  3).

— Методология Kaizen - помогает командам улучшать свою эффективность и качество работы, используя небольшие, но постоянные шаги в развитии процессов.

— Математические формулы для анализа данных, такие как корреляция, регрессия и анализ вариации, могут помочь определить связи между различными факторами и оценить, как они влияют на качество работы команды.

— Методология Lean Six Sigma - комбинация подходов Lean и Six Sigma, которая специализируется на повышении качества работы команды и оптимизации процессов.

### **Заключение.**

Веб-приложение по контролю выполненной работы в команде является важным инструментом управления проектами, который позволяет следить за ходом работы и оценивать эффективность использования времени и ресурсов.

Разработанное приложение использует современные технологии анализа данных и визуализации информации, чтобы обеспечить более точный и полный анализ продуктивности команды.

Такое приложение помогает улучшить управление проектами, экономить время и повышать качество работы. Они также позволяют анализировать данные и принимать корректирующие меры в случае, если сроки реализации сбиваются с графика.

В целом, веб-приложение по контролю выполненной работы в команде представляет собой эффективный и важный инструмент для тех, кто занят в сфере управления и разработки проектов.

### **Список литературы**

[1] Global Workplace Analytics [Электронный ресурс]. URL: <https://globalworkplaceanalytics.com/> (дата обращения 10.04.2023)

[2] Исследование рынка труда [Электронный ресурс]. URL: <https://hh.ru/article/research> (дата обращения 10.04.2023).

[3] Статистика удаленной работы в мире в 2023 году [Электронный ресурс]. URL: <https://inclient.ru/remote-work-stats/> (дата обращения 10.04.2023).

[4] Сколько людей работает удаленно в различных отраслях на 2022 год. [Электронный ресурс]. URL: <https://devby.io/news/skolko-ludei-rabotaet-na-udalonyonke-v-raznyh-otraslyah> (дата обращения 10.04.2023).

## **BIG DATA IN CROWDFUNDING AS A MEANS OF PREDICTING THE SUCCESS OF A CAMPAIGN**

***P.A. Takunov***

*4th grade student of the Faculty of  
Computer Systems and Networks,  
specialty CSaPT BSUIR*

***S.N. Nesterenkov***

*Dean of the Faculty of Computer  
Systems and Networks of BSUIR,  
PhD of Technical Sciences, Associate  
Professor*

***A.N. Markov***

*Senior lecture of the Faculty of  
Computer Systems and Networks of  
BSUIR*

*Department of Computer Science*

*Faculty of Computer Systems and Networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: pavel.takunov@gmail.com*

**Abstract.** The web application for monitoring the work done in a team is a highly effective project management tool that allows you to monitor the tasks performed, evaluate the efficiency of using time and resources, as well as coordinate the work of the entire team.

It allows you to create projects, assign tasks to them and divide them among team members. Users can monitor the status of tasks, track progress and deadlines, as well as share comments and files.

The application is equipped with the functions of analyzing the productivity of the team, as well as the ability to make adjustments to the project if the implementation dates go off schedule.

**Keywords:** web application, performance monitoring, teamwork, project management, productivity analysis, deadlines, data analysis, information visualization, Big Data.

УДК 004.42:004.9

## ИСПОЛЬЗОВАНИЕ ПРОГРАММНО-ИНФОРМАЦИОННОГО КОМПЛЕКСА АРИОН В ИТ-ОБРАЗОВАТЕЛЬНОЙ СРЕДЕ



**С.М. Боровиков**

Доцент кафедры проектирования информационно-компьютерных систем БГУИР, кандидат технических наук, доцент  
*bsm@bsuir.by*



**С.К. Дик**

Доцент кафедры электронной техники и технологии БГУИР, кандидат физико-математических наук, доцент  
*sdick@bsuir.by*

### **С.М. Боровиков**

Доцент кафедры проектирования информационно-компьютерных систем БГУИР. Основная область научных интересов – прикладные математические методы в проектировании изделий радиоэлектроники, включая алгоритмы статистического прогнозирования надёжности изделий электронной техники и оценку надёжности прикладного программного обеспечения на ранних этапах его разработки.

### **С. К. Дик**

Окончил Минский радиотехнический институт по специальности «Радиотехника», руководит научными исследованиями в области лазерной медицины и биомедицинской оптики.

**Аннотация.** Разработанный программно-информационный комплекс по автоматизированному расчёту и обеспечению надёжности электронных устройств и систем, получивший название система АРИОН, был разработан для решения промышленных задач. Особенностью системы является наличие в ней мощной базы данных об элементах отечественного и иностранного производства. Приводятся сведения об интегрировании этой системы в ИТ-образовательную среду для подготовки студентов радиоэлектронных специальностей.

**Ключевые слова:** программно-информационный комплекс, базы данных, автоматизированный расчёт надёжности электронных устройств, ИТ-образовательная среда.

### **Введение.**

В настоящее время использование ИТ-образовательных сред можно рассматривать в качестве эффективного способа как переподготовки специалистов разного профиля, так и обучения студентов технических специальностей. ИТ-образовательные среды могут успешно использоваться при подготовке студентов очной и заочной форм обучения. Чтобы подготовка студентов была эффективной, ИТ-образовательные среды необходимо наполнять современным содержанием. В качестве примера решения таких задач рассматривается включение в ИТ-образовательную среду программно-информационного комплекса (названного системой АРИОН), предназначенного для оценки показателей безотказности электронных устройств и систем разного функционального назначения. Название «система АРИОН» представляет аббревиатуру слов «система Автоматизированного Расчёта и Обеспечения Надёжности». Система была разработана в Белорусском государственном университете информатики и радиоэлектроники по заказу Министерства промышленности Республики Беларусь и может рассматриваться как белорусский вариант подобных российских систем АСОНИКА-К (после переименована в АСОНИКА-Б), АСРН, зарубежных систем RELEX®, ReliaSoft Office Lambda Predict®, RAM Commander и др., ориентированных на автоматизированные методы оценки и анализа надёжности технических устройств [1]. Система АРИОН представляет собой модульный программный комплекс для ПЭВМ, работающий под управлением любой версии операционной

системы Windows, начиная с Windows 2000. Система АРИОН имеет некоторые функции, не реализованные в подобных зарубежных системах, что позволяет в интерактивном режиме работы пользователя с ПЭВМ решать некоторые специфические задачи, в том числе, связанные с прогнозированием индивидуальной надёжности изделий электронной техники.

#### **Актуальность создания и использования системы АРИОН в образовании.**

Окончательный (уточнённый) расчёт показателей надёжности сложных электронных устройств и технических систем является трудоёмким для специалистов по отношению ко всей работе над разрабатываемыми проектами. Он предполагает большое количество математических операций, а результаты расчётов и их достоверность зависят от большого числа конструкторско-технологических, эксплуатационных и других факторов, которые постоянно уточняются при доработке инженером проектного решения. Изменение на этапе проектирования даже одного из этих факторов влечёт за собой необходимость выполнения расчётов заново, что повышает риск допустить неточность и, следовательно, сделать ошибочные выводы. Кроме того, поиск с использованием технической документации сведений об элементах электронных устройств, отвечающих требованиям проектируемой конструкции, занимает у инженера много времени. Для устранения этой проблемы требуется создание мощной базы данных, включающей сведения как об отечественных элементах (страны СНГ), так и об элементах иностранного производства. Причём о каждом элементе электронного устройства необходимо иметь расширенные сведения как об его электрических функциональных параметрах, так и конструкторско-технологических особенностях, и эксплуатационных свойствах. Что касается получения студентами радиоэлектронных специальностей умений по оценке надёжности радиоэлектронной аппаратуры, то во многих случаях поиск необходимой информации об элементах электронных устройств и систем вызывал у обучающихся определённые затруднения. В ряде случаев студентам не удавалось найти нужную информацию, что вынуждало некоторых из них прибегать к искажению исходных данных и итоговых результатов оценки эксплуатационной надёжности электронных устройств и систем. Поэтому создание автоматизированной системы по расчёту надёжности электронного оборудования с мощной базой данных об элементах отечественного и иностранного производства было актуальным и своевременным.

#### **Особенность использования системы АРИОН в образовании.**

Вначале разработанная система АРИОН была ориентирована на проектные организации и производственные предприятия и вызвала интерес у специалистов [2–4], а затем она была адаптирована для использования в учебном процессе, в первую очередь в курсовом и дипломном проектировании студентов радиоэлектронного профиля. Система АРИОН вызвала заметный интерес на республиканских научно-методических конференциях и выставках [5–7].

В учебном процессе показатели надёжности проектируемых электронных устройств студентам предлагается получать с помощью системы АРИОН, включённой в ИТ-образовательную среду по некоторым специальностям радиоэлектронного профиля. Система АРИОН позволяет:

- выполнять автоматизированную оценку (прогнозирование) показателей надёжности электронных устройств и систем на этапе их проектирования;
- производить целенаправленные действия по обеспечению заданных показателей надёжности электронного устройства, выбирая из базы данных элементы с лучшими справочными показателями безотказности;
- выполнять поиск элементов, отвечающих требованиям по эксплуатационным показателям;
- анализировать вклад того или иного элемента в общую ненадёжность электронного устройства.

Система АРИОН проста в использовании, сконструирована так, что сама процедура автоматизированного выполнения расчётов надёжности электронного устройства не снижает понимание студентами сути самих инженерных расчётов.

В системе АРИОН практически все поправочные коэффициенты, используемые в моделях прогнозирования эксплуатационной надёжности элементов, чётко привязаны к определённым факторам, выбор значений которых осуществляется предельно понятным способом из «выпадающих списков» (рисунок 1). При необходимости информацию об элементах можно получать из базы данных, либо вводить значения коэффициентов вручную, что делает гибким процесс обеспечения требования к эксплуатационной надёжности устройств и систем.

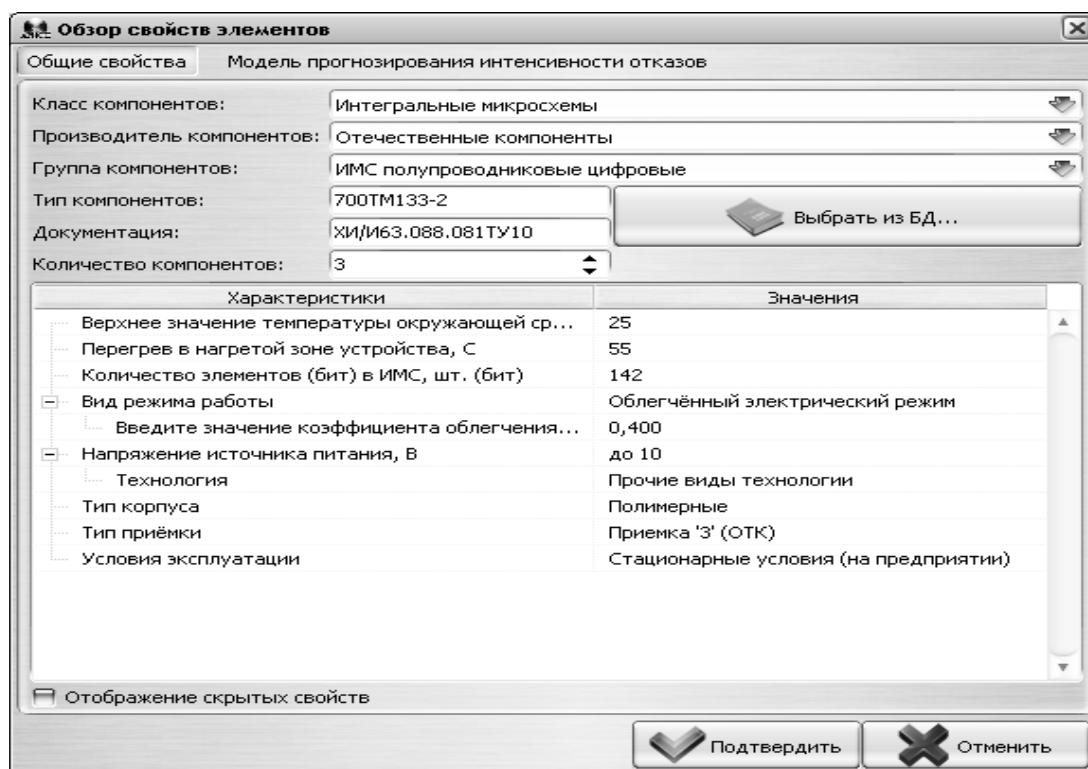


Рисунок 1. Окно обзора свойств компонентов

Отличительной особенностью системы АРИОН является простота интерфейса, что делает систему легко осваиваемой и удобной в IT-образовательной среде. Наглядность представления данных даёт возможность оценить уровень эксплуатационной надёжности не только всего электронного устройства, но и каждого элемента в отдельности.

Результаты автоматизированного расчёта могут быть представлены в следующем виде:

- протокола расчёта (выводится информация об эксплуатационной интенсивности отказов электронного устройства и модулей в его составе);
- столбиковой диаграммы, показывающей вклад каждой части (элемента, модуля) в ненадёжность электронного устройства в целом (рисунок 2);
- документа в формате HTML.



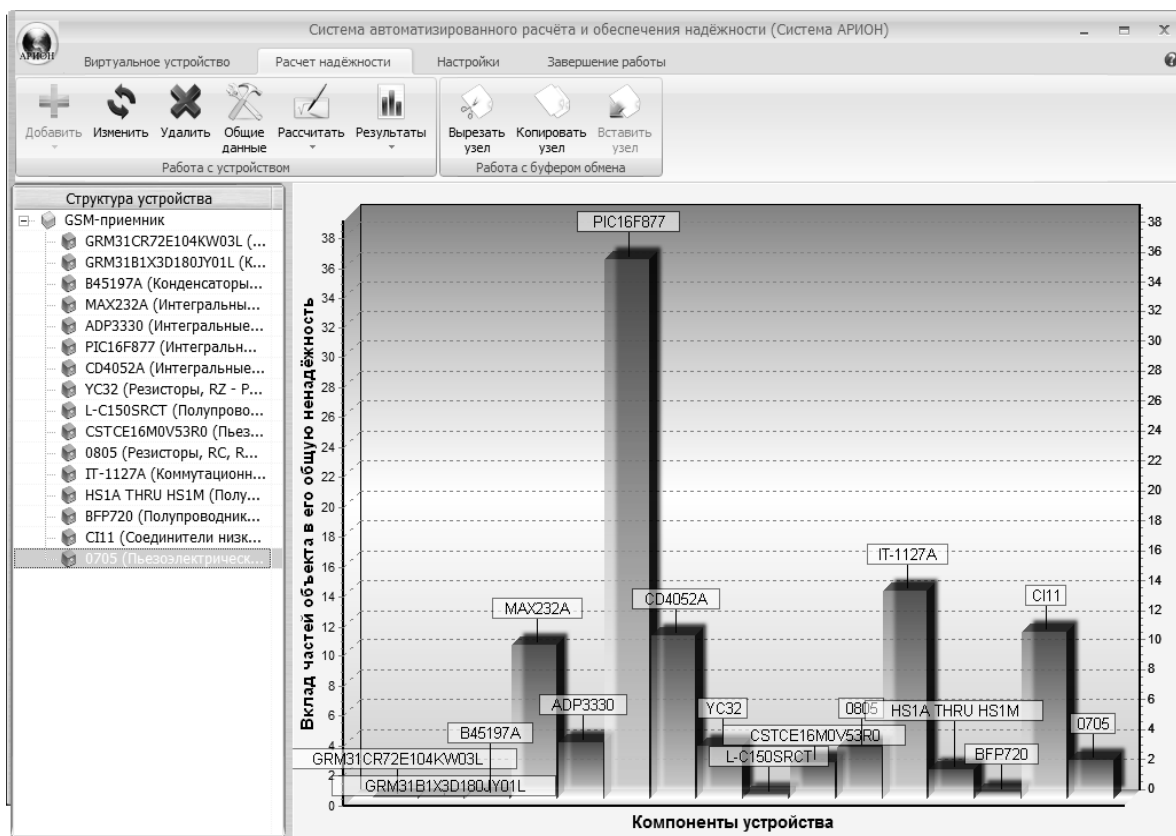


Рисунок 2. Графическое отображение вклада элементов в ненадёжность электронного устройства

Отчёт в формате HTML содержит наиболее полную информацию о результатах расчёта и может рассматриваться как основной. В нём приводится следующая информация:

- общие исходные данные и информация об используемых элементах отечественного и иностранного производства;
- поправочные коэффициенты и значения эксплуатационной интенсивности отказов электронного устройства и его модулей (при их наличии);
- диаграмма вклада составных компонентов (элементов и модулей) в ненадёжность электронного устройства в целом;
- количественные показатели надёжности электронного устройства и эксплуатационной надёжности каждого элемента.

#### **Модернизация программного комплекса АРИОН.**

В 2019 году в программный комплекс АРИОН были добавлены модули [8], позволяющие решать задачи прогнозирования надёжности полупроводниковых приборов и интегральных микросхем на основе моделей, получаемых с помощью методов, описанных в [9, 10]. Использование методов предполагает проведение обучающих экспериментов, результатами которых являются большие массивы данных, вводимые в соответствующие модули модернизированной системы АРИОН+.

На рисунке 3 показаны в виде каскада главное окно и окно обзора компонентов (элементов) программного комплекса «Система АРИОН+».

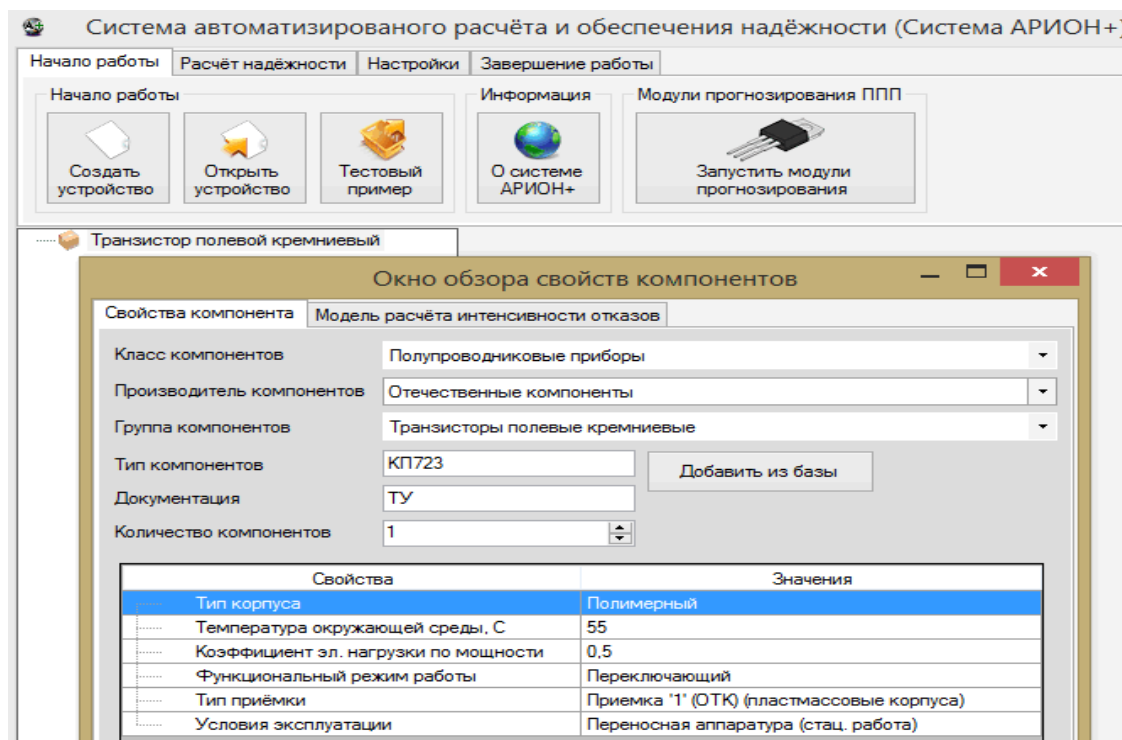


Рисунок 3. Вид главного окна модернизированного программного комплекса АРИОН+

Полученные на основе обработки результатов обучающих экспериментов модели прогнозирования надёжности изделий электронной техники позволяют выполнять:

- индивидуальное прогнозирование надёжности полупроводниковых приборов и интегральных микросхем на основе контроля их информативных параметров в начальный момент времени;
- индивидуальное прогнозирование надёжности полупроводниковых приборов по постепенным отказам на основе имитационной модели с использованием электрического тока или электрического напряжения в качестве имитационного воздействия.

### Заключение.

Внедрение системы АРИОН в IT-образовательную среду повысит точность и достоверность выполняемых студентами инженерных расчётов в курсовых и дипломных проектах, обеспечит более качественную подготовку обучающихся в области надёжности электронного оборудования. Включение в программно-информационный комплекс дополнительных модулей, ориентированных на решение задач прогнозирования надёжности полупроводниковых приборов и интегральных микросхем, расширит возможности использования студентами системы АРИОН в составе IT-образовательной среды.

### Список литературы

- [1]. Разработать систему автоматизированного расчёта показателей надёжности электронных устройств: отчёт о НИР (заключительный) / Белорусский государственный университет информатики и радиоэлектроники; рук. С.М. Боровиков; исполн.: С.М. Боровиков [и др.]. – Минск, 2009. – 146 с. (№ ГР 200.90.344).
- [2]. Управление качеством и надёжностью электронных устройств в системе АРИОН / С.М. Боровиков [и др.] // Информационные технологии, электронные приборы и системы ITEDS` 2010: материалы Междунар. науч.-практ. конф., 6–7 апреля 2010 г., Минск / Белорусский государственный университет. – Минск: Национальная библиотека Беларуси, 2010. – С. 175–177.
- [3]. Боровиков, С.М. IT-комплекс автоматизированного расчёта эксплуатационной надёжности элементов и электронных устройств / С.М. Боровиков // Информационные технологии и системы 2013 г. (ITS-2013): материалы

Международной научной конференции, БГУИР, Минск, Беларусь, 23 октября 2013. – Минск: БГУИР, 2013. – С. 248–249.

[4]. Система автоматизированного расчёта показателей надёжности электронных устройств / С.М. Боровиков [и др.] // Приборостроение–2011: Материалы 4-й Международной НТК, 16–18 ноября 2011 г., Минск, Республика Беларусь. – Минск: БНТУ, 2011. – С. 35–36.

[5]. Боровиков, С. М. Промышленная система АРИОН в обеспечении инженерной подготовки педагогов-радиоинженеров / С.М. Боровиков, О.С. Лосик, Е.Н. Шнейдеров // Инженерно-педагогическое образование: проблемы и пути развития: материалы Международной научно-практической конференции, Минск, 19–20 мая 2011. В 2-х частях. Ч. 2. – Минск: МГВРК, 2011. – С. 7–9.

[6]. Боровиков, С.М. Расчёт надёжности электронных устройств в курсовом и дипломном проектировании с помощью системы АРИОН / С.М. Боровиков, И.Н. Цырельчук, Е.Н.Шнейдеров // Непрерывное профессиональное образование: состояние и перспективы развития: тез. докл. науч.-метод. конф., Минск, 8–9 сентября 2011 г. – Минск: БГУИР, 2011. – С. 34–36.

[7]. Применение системы АРИОН в IT-образовательных средах / С.М. Боровиков [и др.] // Дистанционное обучение – образовательная среда XXI века: матер. VII Международ. научно-метод. конф. (Минск, 1–2 декабря 2011 года). – Минск: БГУИР, 2011. – С. 483–485.

[8]. Разработка программного комплекса автоматизированной оценки надёжности электронных устройств и систем: отчёт о НИР (заключительный) / Белорусский государственный университет информатики и радиоэлектроники; рук. С.М. Боровиков; исполн.: С.М. Боровиков [и др.]. – Минск, 2016. – 45 с. (№ ГР 20121425).

[9]. Боровиков, С.М. Статистическое прогнозирование для отбраковки потенциально ненадёжных изделий электронной техники: монография / С.М. Боровиков. – М.: Новое знание, 2013. – 343 с.

[10]. Прогнозирование надёжности изделий электронной техники / С.М. Боровиков [и др.]; под ред. С.М. Боровикова. – Минск: МГВРК, 2010. – 308 с.

## **USING THE ARION SOFTWARE AND INFORMATION COMPLEX AS AN IT-EDUCATIONAL ENVIRONMENT**

***S.M. Borovikov***

*Associate Professor of the Department of Information  
Computer Systems Design, PhD of Technical  
sciences, Associate Professor*

***S.K. Dick***

*Associate Professor of the Department of Electronic  
Engineering and Technology, PhD of Physical and  
Mathematical Sciences, Associate Professor*

*Department of Information and Computer Systems Design*

*Faculty of Computer Engineering*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: bsm@bsuir.by*

**Abstract.** To solve industrial problems, a software and information complex was developed to perform automated calculations and ensure the reliability of electronic devices and systems, called the ARION system. A feature of the system is the presence in it of a powerful database of elements of domestic and foreign production. Information is given on the integration of this system into the IT educational environment for the preparation of students of radio electronic specialties.

**Keywords:** software and information complex, databases, automated calculation of the reliability of electronic devices, IT-educational environment.

УДК 004.9-027.45

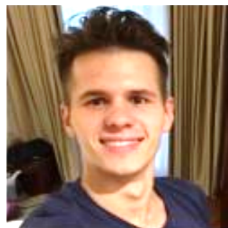
## РЕКОМЕНДАЦИИ ПО ОЦЕНКЕ И ОБЕСПЕЧЕНИЮ НАДЁЖНОСТИ ПЛАНИРУЕМЫХ К РАЗРАБОТКЕ ПРИКЛАДНЫХ КОМПЬЮТЕРНЫХ ПРОГРАММ ДЛЯ ИНФОРМАЦИОННЫХ СИСТЕМ



**Ван Там Лэ**  
Ассистент-исследователь  
Ханойского технологического института, Вьетнам,  
магистр техники и технологии  
luct.tle94@gmail.com



**С.М. Боровиков**  
Доцент кафедры проектирования информационно-компьютерных систем  
БГУИР, кандидат технических наук,  
доцент  
bsm@bsuir.by



**С.С. Дик**  
Проект-менеджер  
компании «Itransition»,  
магистр техники и технологии  
sdick@bsuir.by



**А.В. Будник**  
декан факультета инженеринга и технологий БГАС,  
кандидат технических наук,  
доцент  
A.Budnik@bsac.by

### **Ван Там Лэ**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов: информационные технологии; исследование, разработка и технология изготовления интегральных схем; исследования и разработка технологии изготовления магнитомягких материалов и оптического стекла.

### **С.М. Боровиков**

Окончил Минский радиотехнический институт. Основная область научных интересов – прикладные математические методы в проектировании изделий радиоэлектроники, включая алгоритмы статистического прогнозирования надёжности изделий электронной техники и оценку надёжности прикладного программного обеспечения на ранних этапах его разработки.

### **С.С. Дик**

Окончил Белорусский государственный университет информатики и радиоэлектроники (2016 г.), и аспирантуру при этом университете (2020 г.). Работает в компании «Itransition» в должности проект-менеджера, занимается разработкой и внедрением программного обеспечения в различные сферы деятельности людей.

### **А.В. Будник**

Окончил Минский радиотехнический институт. Основные направления научной деятельности – микроэлектроника, защита информации в технических системах.

**Аннотация.** Систематизированы результаты исследований по оценке надёжности разрабатываемых компьютерных программ, предназначенных для использования в различных сферах деятельности людей.

Приводятся рекомендации по оценке и обеспечению эксплуатационной надёжности планируемых к разработке прикладных компьютерных программ для сложных информационных систем, в том числе используемых для обработки больших данных.

**Ключевые слова:** информационные системы, прикладные компьютерные программы, обеспечение надёжности, тестирование.

### **Введение.**

Важной задачей технологии обработки больших данных (Big Data) является получение аналитических отчётов и принятие на их основе прогнозных решений, которые будут приняты во внимание и использованы компаниями и организациями в своей деятельности. В деятельности компаний

и организаций по изготовлению продукции, оказанию коммерческих, образовательных, медицинских и др. услуг можно выделить важнейшую целевую переменную. В задачах прогнозирования с использованием больших данных целевая переменная может предсказываться на основе большого набора признаков, большая часть из которых может носить косвенный характер относительно процесса или объекта, к которому относится интересующая целевая переменная. Следует стремиться, чтобы набор признаков был достаточным для прогнозирования целевой переменной с требуемой достоверностью [1]. Результаты прогноза обеспечат оперативность принятия решения в определённой области деятельности людей. После выделения целевой функции и определения возможного набора признаков следует определиться с алгоритмами их обработки. Это непростая и очень важная задача. После решения этой задачи возникает очередная, не менее важная задача, состоящая в разработке программного средства по обработке данных. Основу программного средства составляет компьютерная программа. Для обеспечения эффективности обработки данных и получения интересующих прогнозных показателей необходимо, чтобы компьютерная программа отвечала заданным требованиям по её надёжности.

#### **Актуальность.**

По оценке специалистов [2, 3], в сложных информационных системах, к которым относятся компьютерные информационные системы, в том числе использующие большие данные (Big Data), вклад программного обеспечения в ненадёжность систем может составлять 40 и более процентов. Компьютерные программы, разрабатываемые для подобных информационных систем, могут содержать сотни тысяч, и даже миллионы строк кода. После написания программного кода и устранения явных ошибок, обусловленных нарушением грамматики языка программирования (синтаксических ошибок), программа характеризуется начальным уровнем надёжности, который обычно не отвечает требованиям пользователя. Особенностью компьютерных программ, используемых в технологиях Big Data, является высокая изменчивость исходных данных, причём некоторые из них могут носить неструктурированный характер. Этот факт накладывает дополнительные требования к надёжности компьютерных программ. Известны десятки моделей надёжности программных средств, разработанные ранее [4, 5], но вопрос о том, как количественно оценить надёжность компьютерных программ, остаётся в основном проблемным и нерешённым. Каждая организация, являющаяся разработчиком программного средства, или компания, заинтересованная в количественной оценке надёжности программного средства, фактически вынуждена решать проблему для себя заново, учитывая опыт и квалификацию своих программистов и тестировщиков, а также особенность проектируемой информационной системы, использующей большие объёмы сведений, в определённой степени являющимися неструктурированными данными.

Ограничением большинства разработанных ранее методов и моделей оценки надёжности прикладных компьютерных программ является то, что они предполагают наличие определённых данных о тестировании разрабатываемой компьютерной программы после устранения в ней явных ошибок, вызванных нарушениями правил языка программирования. В большинстве случаев заказчиков программного обеспечения для информационных систем интересует ожидаемый эксплуатационный уровень надёжности используемых прикладных компьютерных программ ещё до написания их программного кода и процесса тестирования. Поэтому актуальной является задача по определению прогнозного уровня эксплуатационной надёжности планируемых к разработке компьютерных программ, включая определение требуемого времени их тестирования. При этом следует помнить, что целью тестирования является не тотальное обнаружение всех ошибок, что принципиально невозможно для сложных компьютерных программ, а выявление наибольшего количества наиболее критичных ошибок с точки зрения выполнения компьютерной программой своих основных функций.

#### **Рекомендации, по оценке эксплуатационной надёжности прикладных компьютерных программ.**

На основе анализа ранее выполненных работ, результаты которых описаны в [6–17], в данной статье приводятся некоторые рекомендации, которые могут оказаться полезными для специалистов, занимающихся оценкой и обеспечением эксплуатационной надёжности планируемых к разработке прикладных компьютерных программ для информационных систем, в том числе и ориентированных на

обработку больших данных, характеризующихся значительной изменчивостью. Ниже приводятся основные рекомендации:

1. Важным этапом для планируемой к разработке компьютерной программы является прогнозная оценка количества строк программного кода, в англоязычном варианте – Line Of Code (LOC). Эта единица измерения, по мнению многих специалистов, позволяет установить соотношение с числом ошибок, оставшихся в компьютерной программе. В некоторых кругах специалистов по программированию число строк кода стало единственным приемлемым средством измерения объема компьютерных программ [2]. В случае затруднений в определении прогнозного значения LOC разрабатываемой компьютерной программы следует обратиться к документу [18].

2. В качестве основной характеристики безотказности выполнения программой своих функций, согласно [19], можно использовать вероятность того, что прикладная компьютерная программа безотказно выполнит обработку одного произвольного набора исходных данных из числа тех наборов, которые могут поступать в условиях функционирования компьютерной программы в составе информационной системы. Согласно работе [20] эту вероятность  $P_1$  предлагается определять в предположении экспоненциального закона распределения времени проявления отказов при использовании компьютерной программы:

$$P_1 = \exp(-\lambda_{\text{экс}} \cdot t_1), \quad (1)$$

где  $\lambda_{\text{экс}}$  – ожидаемая (прогнозируемая) интенсивность отказов компьютерной программы, размерность  $1/\text{ч} = \text{ч}^{-1}$ ;  $t_1$  – среднее время (в часах) обработки прикладной компьютерной программой одного набора исходных данных.

В формуле (1) в качестве времени  $t_1$  следует рассматривать процессорное время обработки данных (без учёта времени ожидания).

Факторы, оказывающие наибольшее влияние на значение  $\lambda_{\text{экс}}$  и, следовательно, вероятность  $P_1$ :

- прогнозное число оставшихся в прикладной компьютерной программе скрытых ошибок и их размещение в структуре прикладной компьютерной программы;
- степень изменчивости потока наборов исходных данных на входе прикладной компьютерной программы;
- уровень нагрузки на прикладную компьютерную программу со стороны эксплуатационной среды (загрузка и выгрузка модулей прикладной компьютерной программы из памяти, нахождение операций ввода-вывода в очереди, наличие состояний ожидания и т.д.);

Что касается влияния быстродействия процессора компьютера, то на компьютерах с более высоким быстродействием процессора интенсивность проявления ошибок будет иметь большее значение. В то же время в этом случае процессорное время обработки данных будет меньше. В общем случае произведение  $\lambda_{\text{экс}} \cdot t_1 = \text{const}$ , т. е. быстродействие процессора компьютера фактически не влияет на значение вероятности, определяемой по формуле (1).

3. Важным вопросом является уточнение частоты обращения к прикладной компьютерной программе, т. е. среднего числа сеансов («прогонов») программы за один час в процессе функционирования информационной системы (обозначим эту величину через  $\eta$ ). Тогда вероятность того, что оставшиеся ошибки в прикладной компьютерной программе не проявятся в течение заданного календарного времени  $\tau$ , можно определить, используя выражение:

$$P(\tau) = (P_1)^{\eta \cdot \tau}. \quad (2)$$

Если проектное значение вероятности  $P(\tau)$  не отвечает требованиям функционирования информационной системы, то следует увеличить проектное время тестирования прикладной компьютерной программы.

4. Для оценки прогнозной эксплуатационной интенсивности проявления скрытых ошибок прикладной компьютерной программы (обозначим через  $\lambda_{\text{экс}}$ ) с учётом факторов, указанных в п. 2, приемлемо использование модели, приводимой в [13]:

$$\lambda_{\text{экс}}^{(i)} = \frac{60}{Q} K_{\Sigma}^{(i)} V_{\text{пик}} F_0^{(i)} L \cdot 10^{-6}, \text{ч}^{-1}, \quad (3)$$

где верхний индекс  $i$  указывает на то, что соответствующие характеристики относятся к прикладной компьютерной программе  $i$ -й области применения информационной системы.

Пояснение параметров модели (3) и величин, определяющих эти параметры, приводится в таблице 1.

Таблица 1. Пояснение параметров модели (3)

Обозначение параметра	Пояснение
$Q$	Коэффициент эффективности тестирования, показывающей во сколько раз уменьшается интенсивность отказов компьютерной программы после выполнения её тестирования ( $\lambda_{\text{экс}}$ ) относительно начальной интенсивности отказов $\lambda_0$
$K_{\Sigma}$	Коэффициент увеличения интенсивности отказов из-за суммарного действия изменчивости входных данных и рабочей нагрузки на компьютерную программу со стороны факторов эксплуатационной среды (ввод данных, использование принтера, ожидание операций в очереди и т. д.)
$V_{\text{пик}}$	Пиковая скорость выполнения команд компьютерной программы (скорость для пикового быстродействия процессора $R$ ), определяемая отношением $V_{\text{пик}} = R/B$
$R$	Пиковое быстродействие процессора, указываемое в технической документации; размерность «операций в секунду»
$B$	Прогнозная оценка количества выполняемых процессором команд прикладной компьютерной программы при обработке экстремального потока данных
$F_0$ , ошибка/KLOC	Прогнозная начальная плотность ошибок в компьютерной программе, представляющая собой среднее число ошибок, приходящееся на тысячу строк программного кода (англоязычное сокращение KLOC)
$L$ , KLOC	Прогнозное значение объёма компьютерной программы в KLOC
$F_0 \cdot L = N_0$	Прогнозное значение числа скрытых ошибок в компьютерной программе до выполнения её тестирования

5. Значение параметра  $F_0$  и коэффициента  $K_{\Sigma}$  формулы (3) рекомендуется выбирать по модели RL-92-52 [5], которая с учётом [7–17] принимает вид:

$$F_0 = A \cdot K_{\text{орг}} \cdot K_{\text{кв.прог}} \cdot K_{\text{слож}} \cdot K_{\text{С.Р}} \cdot K_{\text{нов}} \cdot K_{\text{мод}} \cdot \quad (4)$$

В таблице 2 приводится пояснение параметров модели (4), их рекомендуемые значения, а также литературные источники выбора значений применительно к прикладным компьютерным программам, используемым в информационных системах технологий Big Data.



Таблица 2. Выбор параметров модели (4)

Обозначение параметра, размерность	Пояснение	Значение	Литературные источники выбора, примечание
$A$ , ошибка / KLOC	Средняя или базовая плотность ошибок	$\geq 12,8$	$A = 12,8$ ; является максимальным экспериментальным значением [5, 7, 13]
$K_{орг}$	Коэффициент, характеризующий особенность организации, разрабатывающей компьютерную программу	0,5...2,0	[5, 16]
$K_{кв.прог}$	Коэффициент, учитывающий квалификацию и опыт программистов	0,7...2,0	[7, 13, 16]
$K_{слож}$	Коэффициент, учитывающий категорию сложности компьютерной программы	1,0...1,47	[7, 16]
$K_{с.р}$	Коэффициент, учитывающий используемые средства разработки компьютерной программы	0,19...1,3	[7, 13, 16]
$K_{нов}$	Коэффициент, учитывающий степень новизны компьютерной программы	0,63...1,58	[7, 16]
$K_{мод}$	Коэффициент, характеризующий степень использования стандартных модулей в компьютерной программе	0,55...1,0	[7, 13, 16]
$K_{\Sigma}$	Суммарный коэффициент увеличения интенсивности отказов, обусловленный совместным действием изменчивости входных данных и рабочей нагрузки на компьютерную программу	$\geq 19,2$	$K_{\Sigma} = 19,2$ ; является максимальным значением и соответствует автоматизированным системам управления [5, 13]
$r$ , %	Средний процент времени «прогона» компьютерной программы при её тестировании в течение рабочей смены	2,5	[12]

6. Прогнозное количество исполняемых команд (обозначено через  $B$ ) компьютерной программы, используемое для получения прогнозного значения параметра  $V_{пик}$ , подставляемого в формулу (3) можно находить, как:

$$B = L \cdot E_L \cdot E_{ц}, \quad (5)$$

где  $E_L$  – коэффициент расширения кода программы, который определяется языком программирования [16];  $E_{ц}$  – коэффициент увеличения числа выполняемых процессором команд за счёт наличия в программе циклов, условных переходов, ветвлений и других особенностей.

Согласно [21] может быть принято  $E_L = 10$ .

От достоверности прогнозного значения коэффициента  $E_{ц}$  во многом зависит достоверность оценки параметра  $B$  и, следовательно, интенсивности отказов  $\lambda_{экс}$ . Значение  $E_{ц}$  следует получать экспертным методом на основе рассмотрения характера возможных входных данных и анализа алгоритмов обработки этих данных с помощью разрабатываемой компьютерной программы с учётом наихудшего пути её выполнения. Под наихудшим здесь понимается такой путь (трасса процесса обработки данных), для которого количество выполняемых процессором команд прикладной компьютерной программы может принять экстремальное (максимальное) значение. Полученное методом экспертной оценки значение  $E_{ц}$  рекомендуется увеличить на 25–30 процентов.

7. Оценив с использованием произведения (5) прогнозное количество команд  $B$ , выполняемое прикладной компьютерной программой за один прогон в экстремальном режиме, следует найти значение параметра  $V_{пик}$  с учётом быстродействия используемого в компьютере процессора:

$$V_{\text{пик}} = \frac{R}{B},$$

где значение указанного в технической документации быстродействия процессора  $R$  следует подставлять в размерности операций в секунду.

Выбрав по таблице 1 значение коэффициента  $K_{\Sigma}$ , параметр  $A$ , а также значения всех коэффициентов формулы (4), следует по этой формуле определить начальное значение интенсивности отказов ( $\lambda_0$ ) планируемой к разработке компьютерной программы. Для этого в формуле (4) необходимо параметр  $\lambda_{\text{экс}}$  заменить на  $\lambda_0$  и принять  $Q = 1$ .

8. Полученное в п. 7 значение интенсивности отказов  $\lambda_0$ , как правило, не отвечает требованиям для информационной системы, поэтому потребуется процедура тестирования разрабатываемой компьютерной программы в течение определённого времени для обеспечения требуемого (заданного) значения коэффициента эффективности тестирования  $Q_3$ , определяемого как

$$Q_3 = \frac{\lambda_{\text{экс}}^{(3)}}{\lambda_0},$$

где  $\lambda_{\text{экс}}^{(3)}$  – заданная для проектируемой информационной системы эксплуатационная интенсивность отказов компьютерной программы (интенсивность проявления ошибок).

9. Предполагая, что для тестирования будут использоваться типовые методы и технологии тестирования [5, 22], и используя результаты работы [12], прогнозируемое процессорное время тестирования  $t$  можно определить по формуле

$$t \approx \frac{L E_L E_{\Pi} \ln(Q_3)}{60 K_{\Sigma} R}, \quad (6)$$

в которую пиковое быстродействие процессора  $R$  следует подставлять в размерности «млн операций / с», то есть, если  $R = 500 \cdot 10^6$  операций / с, то в качестве  $R$  следует подставлять число 500.

Для быстрого определения влияния параметров модели (6) на значение процессорного времени тестирования  $t$  можно воспользоваться номограммой с двумя бинарными полями, показанной на рисунке 1 [15].

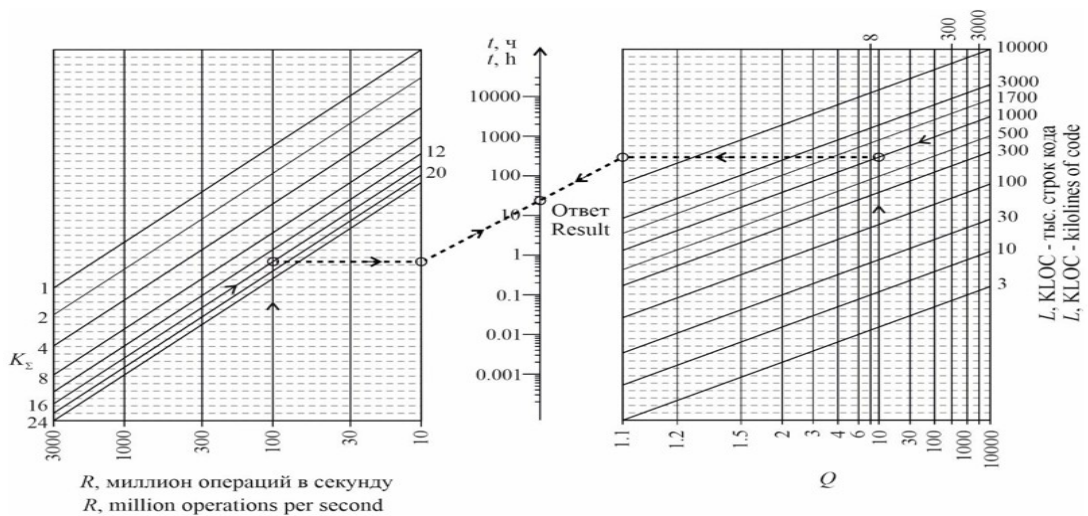


Рисунок 1. Номограмма определения времени тестирования прикладных компьютерных программ

На рисунке 1 показано применение номограммы для следующих значений параметров:  $K_{\Sigma} = 16$ ;  $R = 100$  млн операций / с;  $L = 1000$  KLOC (1000 тыс. строк кода);  $Q_3 = Q = 10$ . Значение  $t$ , полученное по номограмме, примерно равно 23...25 ч и соответствует коэффициентам  $E_L = 1$  и  $E_{ц} = 1$ . В случае использования приведённой номограммы полученное значение  $t$  необходимо умножить на произведение  $E_L \cdot E_{ц}$ .

10. Предполагая, что прикладные компьютерные программы, используемые в технологиях Big Data, наиболее близки к прикладным компьютерным программам, используемым в составе автоматизированных систем управления, выражение, которое рекомендуется для прогнозирования рабочей календарной продолжительности тестирования прикладных компьютерных программ, предназначенных для технологий Big Data, примет вид:

$$T_{\text{календ}}^{(\text{BigData})} = \frac{100t}{r} = \left| r \approx 2,5\% \right| = \frac{100t}{2,5} = 40t.$$

Значение коэффициента  $r$ , показывающего средний процент времени «прогона» компьютерной программы при её тестировании в течение рабочей смены, выбрано по таблице 2. Для прикладных компьютерных программ, используемых в составе автоматизированных систем управления, средний процент выполнения (прогона) компьютерных программ при их тестировании в течение рабочего календарного времени  $T_{\text{календ}}$  составляет примерно 2,5 % [5, 12].

### Заключение.

Приводимые в статье рекомендации позволяют хотя бы примерно оценить ожидаемую надёжность планируемой к разработке прикладной компьютерной программы, используемой в технологии Big Data, и сориентироваться, какое время потребуется для тестирования компьютерной программы с целью обеспечения её эксплуатационной надёжности. Эти оценки важны, поскольку позволяют определить, какие примерно затраты потребуются на разработку прикладной компьютерной программы.

### Список литературы

- [1] Фрэнкс, Б. Укрощение больших данных. Как извлекать знания из массивов / Б. Фрэнкс; пер. с англ. – М.: Изд-во «Технологии развития ООО», 2014. – 352 с.
- [2] Программное обеспечение – источник всех проблем [Электронный ресурс]. – Режим доступа: <http://www.williamspublishing.com/PDF/5-8459-0785-3/part1.pdf> (дата обращения: 17.04.2023).
- [3] Чуканов, В.О. Методы обеспечения аппаратно-программной надёжности вычислительных систем / В.О. Чуканов, В.В. Гуров, Е.В. Прокопьева [Электронный ресурс]. – 2014. – Режим доступа: <http://www.mcst.ru/metody-obespecheniya-apparatnoprogrammnoj-nadezhnosti-vychislitelnykh-sistem>. – Дата доступа: 17.04.2023.
- [4] Sridevi, E. Software reliability – an overview / E. Sridevi, B. Aruna, P. Sowjanya // Intern. J. of Computer Science and Technology. – 2012. – Vol. 3, iss. 1. – P. 471–473.
- [5] McCall, J.A. Software Reliability, Measurement, and Testing Guidebook for Software Reliability Measurement and Testing [Electronic resource] / J.A. McCall [et al.]. – 1992. – Mode of access: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a256164.pdf>. – Date of access: 17.04.2023.
- [6] Боровиков, С.М. Прогнозирование ожидаемой надёжности прикладных программных средств с использованием статистических моделей их безотказности / С.М. Боровиков, С.С. Дик // BIG DATA Advanced Analytics: collection of materials of the fourth international scientific and practical conference, Minsk, Belarus, May 3 – 4, 2018. – Minsk, BSUIR, 2018. – P. 348 – 354.
- [7] Боровиков, С.М. Возможный подход к оценке надёжности прикладных программных средств для технологий Big Data / С.М. Боровиков, Лэ Ван Там, С.С. Дик // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня: сборник материалов V Международной научно-практической конференции, Минск, 13–14 марта 2019 г. В 2 ч. Ч. 2. – Минск: БГУИР, 2019. – С. 77 – 83.
- [8] Модель прогнозирования надёжности планируемых к разработке прикладных компьютерных программ / С.М. Боровиков [и др.] // Интернаука: научный журнал. – 2020. – № 12 (141). – Ч. 1. – С. 68–72.

- [9] Возможный подход к оценке надёжности разрабатываемых программных средств на ранних этапах проектирования информационно-компьютерных систем / С.М. Боровиков [и др.] // Globus: технические науки – от теории к практике [Электронный ресурс]: сборник научных публикаций. – 2020. – Вып. 1 (32). – С. 4–9.
- [10] Боровиков, С.М. Модель прогнозирования ожидаемой надёжности прикладных компьютерных программ / Боровиков С.М., Казючиц В.О. // Информационные радиосистемы и радиотехнологии 2020: материалы Республиканской научно-практической конференции, Минск, 28-29 октября 2020 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск: БГУИР, 2020. – С. 292–295.
- [11] Анализ и оценка надёжности прикладных компьютерных программ / С.М. Боровиков [и др.] // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня: сб. материалов VI Междунар. науч.-практ. конф., Минск, 20-21 мая 2020 года: в 3 ч. Ч. 1. – Минск: Бестпринт, 2020. – С. 382–390.
- [12] Модель прогнозирования времени тестирования прикладных компьютерных программ для технологий BIG DATA / С.М. Боровиков [и др.] // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня: сборник научных статей VII Международной научно-практической конференции, Минск, 19-20 мая 2021 года. – Минск: Бестпринт, 2021. – С. 404–411.
- [13] Borovikov S.M., Kaziuchyts V.O., Khoroshko V.V., Dick S.S., Klinov K.I. Assessment of expected reliability of applied software for computer-based information systems. Informatics, 2021, vol. 18, no. 1, pp. 84–95 (in Russian). <https://doi.org/10.37661/1816-0301-2021-18-1-84-95>.
- [14] Модель прогнозирования времени тестирования прикладных компьютерных программ для автоматизированных систем управления / С.М. Боровиков [и др.] // Информационные технологии и системы 2021 (ИТС 2021) = Information Technologies and Systems 2021 (ITS 2021): материалы международной научной конференции, Минск, 24 ноября 2021 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск: БГУИР, 2021. – С. 28–29.
- [15] Казючиц, В.О. Модель прогнозирования времени тестирования компьютерной программы автоматизированной оценки надёжности полупроводниковых приборов = Model for prediction of testing time of a computer program for automated reliability evaluation of semiconductor devices / В.О. Казючиц, С.М. Боровиков, Е. Н. Шнейдеров // Доклады БГУИР. – 2022. – Т. 20, № 7. – С. 72 – 80. – DOI: <http://dx.doi.org/10.35596/1729-7648-2022-20-7-72-80>.
- [16] Методика обеспечения эксплуатационной надёжности планируемых к разработке прикладных компьютерных программ для информационных систем / С.М. Боровиков [и др.] // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня: сборник научных статей VIII Международной научно-практической конференции, Республика Беларусь, Минск, 11-12 мая 2022 года. – Минск: Бестпринт, 2022. – С. 162–173.
- [17] Лэ, В.Т. Метод оценки надёжности прикладных программных средств на ранних этапах их разработки / В.Т. Лэ, С.С. Дик, С.М. Боровиков // Современные средства связи: материалы XXIII Междунар. науч.-техн. конф., Минск, 18–19 окт. 2018 года. – Минск: БГАС, 2018. – С. 167-168.
- [18] Постановление министерства труда и социальной защиты Республики Беларусь 27 июня 2007 г. № 91 «Об утверждении укрупнённых норм затрат труда на разработку программного обеспечения» [Электронный ресурс]. – Режим доступа: <https://zakonrb.com/npa/ob-utverzhdenii-ukrupnennyh-norm-zatrat-truda> (дата обращения: 17.04.2023).
- [19] Надёжность в технике. Проектная оценка надёжности сложных систем с учётом технического и программного обеспечения и оперативного персонала. Основные положения: ГОСТ 27.205-97. – Введен в действие с 01.10.2005. – Минск: Госстандарт Республики Беларусь, 2005. – 22 с.
- [20] Шубинский, И.Б. Функциональная надёжность информационных систем. Методы анализа / И.Б. Шубинский. – М.: «Журнал Надёжность», 2012. – 296 с.
- [21] Чуканов, В.О. Надёжность программного обеспечения и аппаратных средств систем передачи данных атомных электростанций: учеб. пособие / В.О. Чуканов. – М.: МИФИ, 2008. – 168 с.
- [22] Куликов, С.С. Тестирование программного обеспечения. Базовый курс / С.С. Куликов. – Минск: Четыре четверти, 2017. – 312 с.

## **RECOMMENDATIONS FOR ASSESSING AND ENSURING THE RELIABILITY OF PLANNED FOR DEVELOPMENT APPLIED SOFTWARE FOR INFORMATION SYSTEMS**

***Van Tam Le***

*Research Assistant, Hanoi Institute of Technology, Vietnam, Master of Engineering and Technology*

***S.M. Borovikov***

*Associate Professor of the Department of Information Computer Systems Design, PhD of Technical sciences, Associate Professor*

***S.S. Dick***

*Project manager of the company "Itransition", Master of Engineering and Technology*

***A.V. Budnik***

*Dean of the Faculty of Engineering and Technology of the BSACE, Candidate of Technical Sciences, Associate Professor*

*Department of Information and Computer Systems Design*

*Faculty of Computer Engineering*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: bsm@bsuir.by*

**Abstract.** The results of studies on the reliability assessment of developed software intended for use in various fields of human activity are systematized.

Recommendations are given for assessing and ensuring the operational reliability of the planned development of applied software for complex information systems, including those used for processing big data.

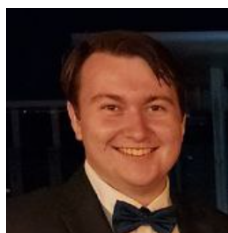
**Keywords:** information systems, application

УДК 004.622

## КАЧЕСТВО ДАННЫХ В BIG DATA



**А.И. Лордкипанидзе**  
Студент 4 курса, кафедры  
ПИКС  
91430072@study.bsuir.by



**В.Д. Владымыцев**  
Ассистент кафедры  
информатики, инженер  
программист ОИАСУ ЦИИР  
БГУИР  
v.vladymtsev@bsuir.by



**А.Н. Марков**  
Старший преподаватель,  
заместитель начальника  
Центра информатизации и  
инновационных разработок  
БГУИР  
a.n.markov@bsuir.by

### **А.И. Лордкипанидзе**

Студент 4 курса “Информационные системы и технологии в бизнес-менеджменте” БГУИР.

### **В.Д. Владымыцев**

Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР

### **А.Н. Марков**

Старший преподаватель кафедры ПОИТ, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Качество данных является одним из ключевых факторов, влияющих на успешность обработки и использование больших объемов данных. В связи с этим вопросы качества данных в big data становятся все более актуальными. В данной аннотации будут рассмотрены основные проблемы, связанные с качеством данных в big data, а также методы и инструменты, используемые для их решения. Особое внимание будет уделено таким аспектам, как точность, полнота, актуальность, надежность, целостность и безопасность данных. Решение проблем, связанных с качеством данных в big data, требует комплексного подхода и использования различных методов и инструментов, включая методы проверки и очистки данных, автоматизированные процессы обновления данных, аутентификацию и контроль доступа, шифрование и многое другое.

**Ключевые слова:** big data, качество данных, точность данных, полнота данных, актуальность данных, надежность данных, целостность данных, безопасность данных, проверка данных, очистка данных, обновление данных, контроль доступа, шифрование данных.

### **Введение.**

В последнее время обработка больших объемов данных, или big data [1], стала неотъемлемой частью работы многих компаний и организаций. Однако, обработка и использование таких больших объемов данных не всегда является простой задачей и может сопровождаться рядом проблем, которые могут повлиять на качество данных и результаты анализа. Одной из таких проблем является качество данных в big data. Качество данных играет важную роль в успешности обработки и использования данных, поэтому данной теме уделяется все больше внимания.

В данном тексте будет рассмотрено, что такое качество данных в big data, какие проблемы связаны с ним, а также какие методы и инструменты используются для обеспечения качества данных.

### **Качество данных в big data.**

Качество данных [2] - одна из ключевых проблем, связанных с обработкой больших объемов данных (Big Data).

Неправильные данные могут привести к неправильным выводам и решениям, поэтому важно обеспечить высокое качество данных.

Таблица 1. Источники данных

Внутренние	Внешние
ERP [3]	Социальные сети
Классификаторы	Интернет
CRM	Специализированные DataSet

Ниже представлены основные из проблем, связанных с качеством данных, а также методы и инструменты, которые могут помочь решить эти проблемы.

#### **Недостаточная точность.**

Данные могут содержать ошибки, неточности и неточности, которые могут привести к неправильным выводам. Одним из способов решения этой проблемы является метод очистки данных, который позволяет идентифицировать и исправить ошибки в данных.

Этот процесс может включать в себя удаление дубликатов, заполнение отсутствующих значений, а также коррекцию ошибок.

#### **Низкая связность данных.**

Данные могут быть разбросаны по разным системам и приложениям, что затрудняет их интеграцию и анализ.

Для решения этой проблемы можно использовать методы интеграции данных, которые позволяют объединить данные из разных источников в единый набор данных. Этот процесс может включать в себя приведение данных к одному формату, обработку дубликатов и определение связей между данными.

#### **Низкая полнота.**

Данные могут быть неполными, что может привести к неправильным выводам и решениям. Одним из способов решения этой проблемы является метод заполнения данных, который позволяет заполнить отсутствующие значения на основе имеющихся данных.

Этот процесс может включать в себя использование методов машинного обучения и статистических методов для предсказания значений.

#### **Низкая консистентность.**

Данные могут быть несогласованными, что может привести к проблемам при интеграции и анализе данных. Для решения этой проблемы можно использовать методы управления данными, которые позволяют обеспечить согласованность данных в разных системах и приложениях.

#### **Низкая актуальность.**

Данные могут устареть, что может привести к неправильным выводам и решениям. Для решения этой проблемы можно использовать методы обновления данных, которые позволяют обновлять данные в реальном времени или периодически, чтобы обеспечить их актуальность.

Для этого могут использоваться различные методы, например, анализ изменений в источниках данных или настройка автоматических процессов обновления данных.

#### **Недостаточная безопасность.**

Обработка больших объемов данных может привести к угрозам безопасности, например, к краже данных, взлому системы или потере данных.

Для обеспечения безопасности данных можно использовать методы шифрования, аутентификации, авторизации и контроля доступа, а также использовать соответствующие инструменты и технологии для обнаружения и предотвращения угроз.

#### **Сложность анализа.**

Обработка больших объемов данных может быть очень сложной и требовать значительных вычислительных ресурсов. Для решения этой проблемы можно использовать



специализированные инструменты и технологии, такие как Hadoop [4], Spark [5] и другие, которые позволяют обрабатывать большие объемы данных быстро и эффективно. Пример использования технологии Hadoop представлен на рисунке 1.

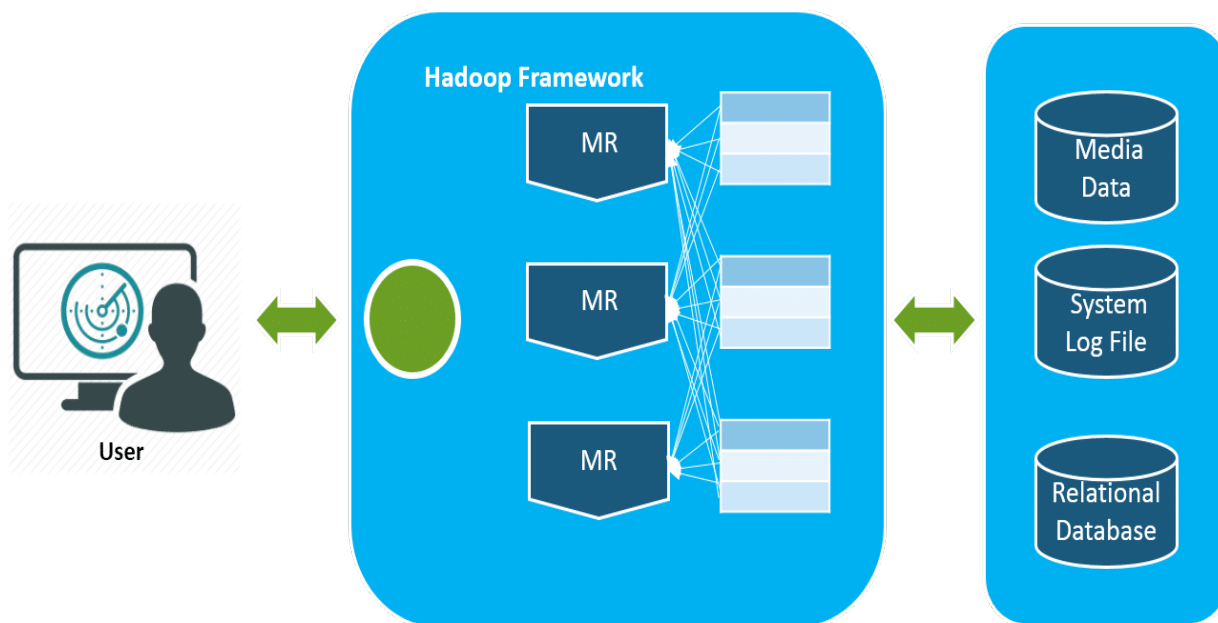


Рисунок 1. Пример использования технологии Hadoop.

#### Недостаточное понимание данных.

Важно иметь хорошее понимание данных, которые вы обрабатываете, чтобы сделать правильные выводы и принять правильные решения. Для этого необходимо проводить анализ данных, чтобы определить их структуру, связи, распределение и другие важные характеристики.

Кроме того, необходимо учитывать контекст, в котором были собраны данные, чтобы правильно интерпретировать их значения и использовать их для принятия решений.

#### Заключение.

Качество данных является важным фактором при обработке больших объемов данных. Проблемы, связанные с качеством данных, могут влиять на результаты анализа и приводить к некорректным выводам.

Однако, современные методы и инструменты позволяют решить многие из этих проблем. Кроме того, улучшение качества данных в big data может повысить эффективность бизнес-процессов и улучшить принятие решений.

Основными методами, используемыми для обеспечения качества данных в big data, являются проверка и очистка данных, обновление данных, контроль доступа, шифрование и многое другое. Однако, необходимо помнить, что обеспечение качества данных является процессом постоянного улучшения, и требует непрерывного мониторинга и анализа.

#### Список литературы

- [1] Теоретический минимум по Big Data. Всё, что нужно знать о больших данных. — СПб.: Питер, 2019. — 208 с.: ил. — (Серия «Библиотека программиста»). ISBN 978-5-4461-1040-7.
- [2] Редман, Т. К. Качество данных: полевое руководство. Издательство Digital Press., 2016 — 260 с.:
- [3] Астапкина, К. С. Анализ и описание современных ERP-систем / Астапкина К. С. // Актуальные научные исследования в современном мире. – 2022. – Вып. 1(81), Ч. 10. – С. 164–166.
- [4] Big Data: Concepts, Technology and Architecture / В. Balusamy [и др.]. – Hoboken : John Wiley & Sons, Inc., 2021. – 368 с

[5] Свирновский, А. В. Ключевые концепции Apache Spark / Свирновский А. В. // Новые информационные технологии в научных исследованиях: материалы XXV Юбилейной Всероссийской научно-технической конференции студентов, молодых ученых и специалистов, Рязань, 18-20 ноября 2020 г. / Рязанский государственный радиотехнический университет им. В. Ф. Уткина. – Рязань : ИП Коняхин А. В. (Book Jet), 2020 – С. 192 – 194.

## **DATA QUALITY IN BIG DATA**

***A.I. Lordkipanidze***

*4th year student, Department of  
PIKS*

***V.D. Vladymtsev***

*Assistant of the Department of  
Computer Science, Software  
Engineer of DIACS CIIR BSUIR*

***A.N. Markov***

*Senior lecturer of the department,  
Deputy head of the Center for  
Informatization and Innovative  
Developments*

*Center for Informatization and Development of the Belarusian University of State Informatics and  
Radioelectronics, Republic of Belarus*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*

*E-mail: 91430072@study.bsuir.by, v.vladymtsev@bsuir.by, a.n.markov@bsuir.by*

**Abstract.** Data quality is one of the key factors that affects the success of processing and using large volumes of data, or big data. Therefore, issues related to data quality in big data are becoming increasingly relevant. This annotation examines the main problems related to data quality in big data, as well as the methods and tools used to address them. Special attention is paid to aspects such as accuracy, completeness, timeliness, reliability, integrity, and security of data. Solving problems related to data quality in big data requires a comprehensive approach and the use of various methods and tools, including data validation and cleansing methods, automated data update processes, authentication and access control, encryption, and much more.

**Keywords:** Data quality, Big data, Accuracy, Completeness, Timeliness, Reliability, Integrity, Security, Data validation, Data cleansing, Automated data update processes, Authentication, Access control, Encryption.

УДК 004.021:004.75

## ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ДАННЫХ В РАЗВИТИИ СОВРЕМЕННОГО ЗЕМЛЕДЕЛИЯ



**С.С. Марковский**  
Студент БГУИР  
stepan.markovskiy@mail.ru



**С.Н. Нестеренков**  
Кандидат технических наук, доцент,  
декан факультета компьютерных  
систем и сетей  
s.nesterenkov@bsuir.by

### **С.Н. Нестеренков**

Кандидат технических наук, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры Программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

### **С.С. Марковский**

Студент 4 курса специальности «Вычислительные машины, системы и сети» БГУИР.

**Аннотация.** Выполнено исследование об использовании и применении больших данных в современном сельском хозяйстве. Изучены основные концепции внедрения больших данных и сферы их применения в сельском хозяйстве. Описаны положительные улучшения на разных этапах выращивания и реализации сельскохозяйственной продукции, достигаемые благодаря внедрению больших данных. Приведена оценка эффективности использования современных технологий, основанных на больших данных, в развитии земледелия.

**Ключевые слова:** точное земледелие, большие данные, вегетационный индекс NDVI, спутниковые изображения, погодные данные, система позиционирования.

### **Введение.**

Питание является жизненно необходимой потребностью для любого человека. Эту потребность люди стремятся удовлетворить каждый день. От правильного питания во многом зависит здоровье человека и его способность справляться с повседневными задачами.

Однако сегодня существуют разные проблемы, из-за которых многие люди по всей земле не могут в полной мере удовлетворить свою потребность в качественной пище.

Современное сельское хозяйство является основным источником продовольствия. От эффективного использования земельных ресурсов зависит объем получаемого продовольствия. Поэтому во многих странах этому сектору уделяется особое внимание.

Существует множество факторов, влияющих на успешное ведение сельского хозяйства. Для более точного анализа этих факторов применяются современные достижения из области больших данных.

Существуют технологии, которые способствуют повышению урожайности и обеспечивают устойчивый рост мирового сельского хозяйства. Если раньше фермеры принимали решения на основании неточных прогнозов и интуиции, то современные инструменты на стыке технологической и сельскохозяйственной отрасли позволят изменить эту ситуацию.

### **Точное земледелие.**

Современное сельское хозяйство сфокусировано на точном земледелии. Точное земледелие – инновационный метод в сельском хозяйстве с использованием новейших технологий для улучшения качества урожая.

Технологии точного земледелия предусматривают использование точных данных дистанционного зондирования, таких как снимки или видеоизображения с дронов или спутников. Такие изображения позволяют эффективно наблюдать за состоянием почв и урожая.

Этот метод земледелия привлекает все больше внимания со стороны аграриев, поскольку помогает сократить расходы и улучшить состояние окружающей среды [1].

Применение метода точного земледелия позволяет достичь следующих преимуществ [2]:

- экономия средств производства;
- экономия рабочего времени персонала и техники;
- полное использование потенциала урожайности;
- повышение качества продукции;
- рост производительности труда;
- рост прибыли за счет повышения урожайности и снижения издержек производства.

#### Большие данные в точном земледелии.

Источники больших данных в точном земледелии — это датчики, которые собирают информацию на земле, и спутниковые снимки.

Сочетание данных из обоих источников позволяет фермерам получить самую точную информацию о состоянии своих полей. Чтобы получить максимальную пользу, для обработки данных нужно использовать машинное обучение и искусственный интеллект.

Есть несколько способов использования больших данных для помощи фермерам. Всё начинается с картирования полей и культур по всему миру. Стартап OneSoil был первым, кто решил эту сложную задачу на мировом уровне. Интерактивная карта OneSoil Map (рисунок 1) предоставляет информацию о полях и посевах в Европе и США, помогает отслеживать тенденции на региональном уровне и следить за изменениями на конкретном поле [3].

Карта работает благодаря обработке спутниковых изображений при помощи алгоритмов машинного обучения.

Анализ спутниковых снимков позволяет следить за состоянием полей удалённо через изменения в вегетационном индексе NDVI. Комбинируя данные, полученные при помощи спутников и датчиков, фермер может рассчитать и применить правильную дозу удобрений и химикатов для каждой части поля.

Еще одна возможность применения больших данных — сбор погодных данных. Сельскохозяйственное производство зависит от погодных условий на протяжении вегетационного периода. Ценной является способность фермера предсказывать экстремальные и неблагоприятные погодные явления и приспосабливаться к ним.

Платформы на базе BigData могут отражать изменения погодных условий в режиме реального времени, чтобы фермеры могли оперативно реагировать на изменения [4].

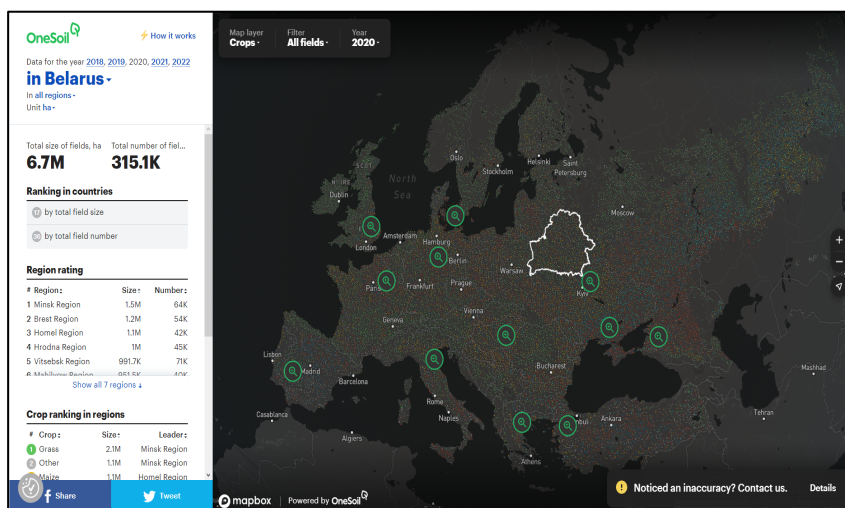


Рисунок 1. Интерактивная карта OneSoil Map

В процессе отслеживания цепи поставок также используются большие данные. Когда мы имеем дело с такими чувствительными материалами как семена, растения и пищевые продукты, предотвращение порчи становится серьезной проблемой.

Большие данные помогают фермерам и поставщикам оптимизировать управление автопарком для повышения надежности доставки.

Интеллектуальные счётчики и системы позиционирования на основе GPS улучшают маршрутизацию, сокращают транспортные расходы и т.д.

#### **Оценка эффективности.**

По одной из оценок, внесение изменений в методы ведения сельского хозяйства и производства продуктов питания, которые повышают производительность, способствуют внедрению методов устойчивого развития экономики и сокращению отходов, может создать коммерческие возможности и новую экономию на общую сумму 2,3 триллиона долларов США в год по всему миру [5].

Благодаря применению концепции точного земледелия происходит оптимизация операционных расходов и повышение урожайности в среднем на 15-20% за счет сокращения объемов используемых семян, агрохимикатов, удобрений и воды, которые применяются строго «по потребности» [6].

#### **Заключение.**

Хотя изначально сельскохозяйственная отрасль не полагалась на новейшие разработки в области компьютерных технологий, их постепенное внедрение в эту отрасль доказывает свой успех.

Использование больших данных в сельском хозяйстве уже сейчас позволяет повысить эффективность производства продукции на многих этапах.

Это может открыть возможность для решения глобальных экологических и продовольственных проблем общества.

#### **Список литературы**

[1] Точное земледелие для эффективного сельского хозяйства: [Электронный ресурс]. URL: <https://eos.com/ru/blog/tochnoe-zemledelie/>. (Дата обращения: 03.04.2023).

[2] Что такое точное земледелие и зачем оно применяется: [Электронный ресурс]. URL: <https://ttz.by/about-technology/>. (Дата обращения: 03.04.2023).

[3] Agricultural OneSoil Map with AI detected fields and crops: [Электронный ресурс]. URL: <https://map.onesoil.ai/2020/BLR#3.37/52.43/24>. (Дата обращения: 04.04.2023).

[4] Что дает сельскому хозяйству использование Big Data: [Электронный ресурс]. URL: <https://aggeek.net/ru/blog/chto-daet-selskomu-hozyajstvu-ispolzovanie-big-data>. (Дата обращения: 04.04.2023).

[5] A game plan for technology companies to actually help save the world: [Электронный ресурс]. URL: <https://theconversation.com/a-game-plan-for-technology-companies-to-actually-help-save-the-world-105007>. (Дата обращения: 05.04.2023).

[6] Digital Agriculture Technology - Adoption & Attitudes Study / Irish Farm Centre, Bluebell, Dublin, 2019. – 68 с.

## **THE USAGE OF BIG DATA IN THE DEVELOPMENT OF MODERN AGRICULTURE**

**S. Markovsky**

*Student of Belarusian State  
University of Informatics  
and Radioelectronics*

**S. Nesterenkov**

*PhD, Associate Professor, Dean of the Faculty  
of Computer Systems and Networks*

*Belarusian State University of Informatics and Radioelectronics  
E-mail: [stepan.markovskiy@mail.ru](mailto:stepan.markovskiy@mail.ru), [s.nesterenkov@bsuir.by](mailto:s.nesterenkov@bsuir.by)*

**Abstract.** The analysis on the use and application of big data in modern agriculture is carried out. The main concepts of big data implementation and their application in agriculture have been studied. Positive improvements at different stages of cultivation and sale of agricultural products, achieved through the implementation of big data, are described. The effectiveness of the usage of modern technologies based on big data in the development of agriculture is assessed.

**Keywords:** precision farming, big data, NDVI vegetation index, satellite images, weather data, positioning system.

УДК 621.865.8:629.055

## К ВОПРОСУ УПРАВЛЕНИЯ МОБИЛЬНЫМ РОБОТОМ С КОЛЕСАМИ ВСЕНАПРАВЛЕННОГО ТИПА



**С.А. Павлюковец**  
Заведующий кафедры  
«Электропривод и  
автоматизация промышленных  
установок и технологических  
комплексов» БНТУ, кандидат  
технических наук, доцент  
sap@bntu.by



**А.А. Вельченко**  
Доцент кафедры  
«Электропривод и  
автоматизация  
промышленных установок и  
технологических  
комплексов» БНТУ,  
кандидат технических  
наук, доцент  
anna.velchenko@gmail.com



**У Синьсинь**  
Аспирант кафедры  
«Электропривод и  
автоматизация промышленных  
установок и технологических  
комплексов» БНТУ  
cotemorning42@gmail.com



**А.А. Радкевич**  
студент кафедры  
«Электропривод и  
автоматизация  
промышленных установок и  
технологических  
комплексов» БНТУ  
artyomradkevichbntu@gmail.  
com



**Н.О. Савко**  
студент кафедры  
«Электропривод и  
автоматизация промышленных  
установок и технологических  
комплексов» БНТУ  
nikitsauko@mail.ru

### **С.А. Павлюковец**

Кандидат технических наук, доцент, заведующий кафедрой «Электропривод и автоматизация промышленных установок и технологических комплексов» Белорусского национального технического университета. Область научных интересов связана с разработкой методов и алгоритмов управления мобильными роботами.

### **А.А. Вельченко**

Кандидат технических наук, доцент, доцент кафедры «Электропривод и автоматизация промышленных установок и технологических комплексов» Белорусского национального технического университета. Область научных интересов связана с разработкой и исследованием нейросетевых регуляторов для электропривода.

### **У Синьсинь**

Аспирант Белорусского национального технического университета. Область научных интересов связана с исследованием алгоритмов управления мобильными роботами с колесами всенаправленного типа.

### **А.А. Радкевич**

Студент 4-го курса Белорусского национального технического университета. Область научных интересов связана с разработкой систем управления электроприводами.



**Н.О. Савко**

Студент 3-го курса Белорусского национального технического университета. Область научных интересов связана с разработкой математических моделей электропривода.

**Аннотация.** В настоящей работе рассмотрены математические модели управления роботами с колесами всенаправленного типа (*omni*-колесо и *mesanum*-колесо), представлены варианты комбинации различных типов всенаправленных колес в работе.

**Ключевые слова:** мобильный робот, всенаправленное колесо, *omni*-колесо, *mesanum*-колесо, управляемость, мобильность.

### **Введение.**

Колесным мобильным роботом (КМР) называется автоматическое транспортное средство, управляемое дистанционно либо имеющее различные степени автономности. Данный фактор обусловил применение КМР для выполнения задач, связанных с риском для здоровья и безопасности человека, а также для автоматизации выполнения однотипных операций в различных отраслях промышленности, на транспорте и т.п.

Классификация КМР осуществляется в соответствии с использованным типом колес. В зависимости от выполняемой функции колеса в мобильных роботах подразделяются на управляемые и неуправляемые, ведущие и ведомые. Под управляемым колесом понимается такое колесо, направление вращения которого может изменяться в результате воздействия управляющей системы. В свою очередь, неуправляемое колесо не способно поворачиваться относительно оси своего вращения. Поворот такого колеса представляется возможным лишь в случае действия на него внешних сил, а не воздействия управляющей системы. Исходя из возможности комбинации различных типов колес, выделяются следующие разновидности колес по отношению к использованию в мобильном роботе:

- роликовые (неуправляемые и ведомые);
- неповоротные и неуправляемые;
- неповоротные и управляемые;
- управляемые и ведущие;
- управляемые и ведомые.

В управляемых колесах ось поворота зачастую пересекает ось вращения колеса, в то время как в неуправляемых эти оси всегда расположены на фиксированном определенном расстоянии относительно друг друга. Кроме того, в КМР могут применяться одинарные и двойные колеса, что не оказывает никакого влияния на кинематическую структуру робота, но влияет на распределение усилий между колесами.

КМР обычно являются неголономными транспортными средствами, движение которых ограничивается их скоростью. На практике это означает, что такие КМР имеют менее трех степеней подвижности. Исключение представляют всенаправленные мобильные роботы, являющиеся голономными транспортными средствами, т.к. они могут реализовать три независимых движения: в продольном и поперечном направлениях, а также вращение вокруг вертикальной оси.

На сегодняшний день вопрос управления колесами всенаправленного типа является одним из актуальных в мобильной робототехнике, поскольку их применение позволяет роботу двигаться во всех направлениях, что в ряде случаев дает выигрыш во времени.

Всенаправленные колеса являются более сложной механической конструкцией по сравнению с другими видами колес, что делает их относительно дорогими. Более того, во всенаправленных роботах все колеса имеют независимый привод и управление, что требует не менее трех электродвигателей.

При управлении мобильным роботом с колесами всенаправленного типа намного сложнее решение задач кинематики и, следовательно, контроля движения мобильной платформы.



Несмотря на это, всенаправленные роботы характеризуются хорошей маневренностью – они могут одновременно вращаться на месте и двигаться в любом направлении.

В процессе движения всенаправленные роботы не обладают хорошей устойчивостью, поскольку их управляемость и маневренность сильно зависит от скорости каждого колеса. Если одно из них утрачивает контакт с землей, робот движется в непреднамеренном направлении. Поэтому их оптимальный режим движения ограничивается ровной поверхностью.

Более того, всенаправленные роботы обладают худшим по сравнению с дифференциально управляемыми роботами навигационным счислением. Поэтому для обеспечения высокой точности движения, управление роботом не должно основываться только на его одометрии.

#### Классификация колес всенаправленного типа.

Разновидностями всенаправленных колес являются и *omni*-колеса и *mecanum*-колеса. Их особенность по сравнению с колесами традиционного типа заключается в том, что во всенаправленных колесах лишь одна компонента скорости точки соприкосновения колеса с плоскостью движения равна нулю при движении.

*Mecanum*-колеса характеризуются тем, что по окружности их обода размещены неприводные ролики, оси вращения которых повернуты на угол  $45^\circ$  относительно оси вращения колеса (рисунок 1). *Omni*-колеса являются двойными колесами, по окружности обода которых также размещены ролики, однако их угол вращения повернут на  $90^\circ$  относительно оси вращения колеса, что изображено на рисунке 2.

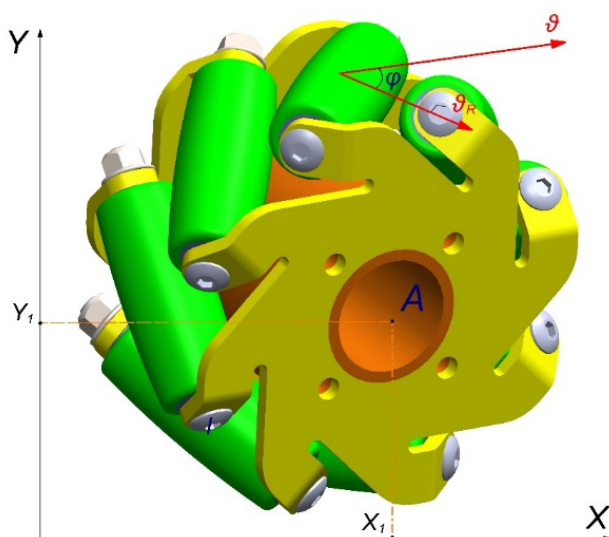


Рисунок 1. Трехмерная математическая модель *mecanum*-колеса

Кинематическая модель колесного робота с *omni*-колесами является голономной системой. В случае использования голономного привода, колеса располагаются таким образом, чтобы оси их вращения пересекались в одной точке. Роботы с голономным приводом обычно имеют три и реже четыре колеса.

Положение *omni*-колеса радиусом  $R$  относительно корпуса робота описывается следующими переменными: начальной точкой  $P$  неподвижной системы координат  $(\vec{X}, \vec{Y})$ , неподвижным центром колеса  $A$  в данной системе координат, положение которого относительно этой системы характеризуется полярными координатами, в частности, длиной радиуса-вектора  $PA = l$  и углом  $\alpha$ . Ориентация плоскости колеса относительно направления радиус-вектора  $PA$  характеризуется постоянным углом  $\beta$ . Угол поворота колеса относительно его горизонтальной оси обозначается как  $\varphi(t)$ .

Размещение *omni*-колеса в неподвижной системе координат относительно корпуса робота, как и в случае с колесом традиционного фиксированного типа, определяется тремя постоянными параметрами:  $\alpha$ ,  $\beta$ ,  $l$ . Для того, чтобы охарактеризовать направление движения, в проекции на которое скорость точки контакта равна нулю, вводится дополнительная переменная  $\gamma$ , являющаяся углом между плоскостью колеса и соответствующим направлением. Тогда соотношение между этими величинами имеет вид:

$$\left[ -\sin(\alpha + \beta + \gamma) \cos(\alpha + \beta + \gamma) l \cdot \cos(\beta + \gamma) \right] \cdot R(\Theta) \cdot \dot{\xi} + r \cdot \cos \gamma \cdot \dot{\varphi} = 0. \quad (1)$$

Кинематическая структура *omni*-колеса с расположением векторов данных переменных и параметров представлена на рисунке 2.

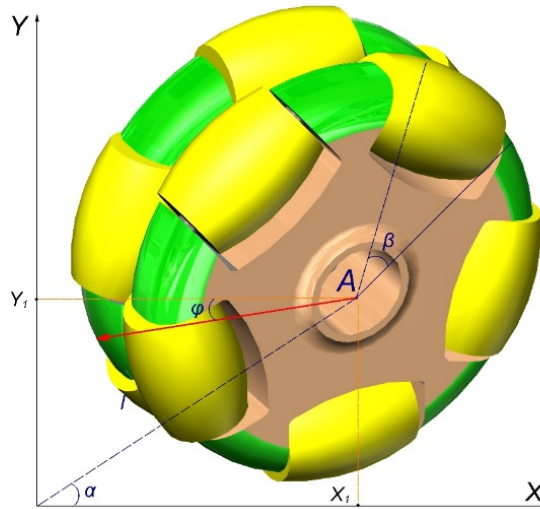


Рисунок 2. Трехмерная математическая модель *omni*-колеса

### Мобильный робот с колесами всенаправленного типа *omni*.

Распространенным примером использования *omni*-колеса в робототехнике является трехколесный мобильный *omni*-робот, кинематическая модель которого представлена на рисунке 3.

Инверсная кинематика такого робота в неподвижных координатах  $(\vec{X}, \vec{Y})$  рассматривается с учетом полной скорости робота  $\mathcal{G}$  и его угловой скорости  $\dot{\varphi}$ :

$$\mathcal{G} = \sqrt{\dot{x}^2 + \dot{y}^2}. \quad (2)$$

Полная скорость первого колеса  $\mathcal{G}_1$  складывается из поступательной  $\mathcal{G}_{1T}$  и вращательной  $\mathcal{G}_{1R}$  составляющих скорости:

$$\mathcal{G}_1 = \mathcal{G}_{1T} + \mathcal{G}_{1R}; \quad (3)$$

$$\mathcal{G}_{1T} = -\dot{x} \cdot \sin(\varphi) + \dot{y} \cdot \cos(\varphi); \quad (4)$$

$$\mathcal{G}_{1R} = R \cdot \dot{\varphi}. \quad (5)$$

Поэтому полная скорость первого колеса в данном роботе равняется:

$$\mathcal{G}_1 = \mathcal{G}_{1T} - \dot{x} \cdot \sin(\varphi) + \dot{y} \cdot \cos(\varphi) + R \cdot \dot{\varphi}. \quad (6)$$

Аналогично, учитывая общий угол поворота второго колеса  $(\varphi + \Theta_2)$ , его полная скорость  $\mathcal{G}_2$  рассчитывается по выражению:

$$\mathcal{G}_2 = -\dot{x} \cdot \sin(\varphi + \Theta_2) + \dot{y} \cdot \cos(\varphi + \Theta_2) + R \cdot \dot{\varphi}. \quad (7)$$

Таким же образом определяется полная скорость третьего колеса с учетом его поворота на угол  $(\varphi + \Theta_3)$ :

$$\mathcal{G}_3 = -\dot{x} \cdot \sin(\varphi + \Theta_3) + \dot{y} \cdot \cos(\varphi + \Theta_3) + R \cdot \dot{\varphi}. \quad (8)$$

Инверсная кинематическая модель трехколесного мобильного *omni*-робота в неподвижной системе координат  $(\vec{X}, \vec{Y})$  определяется матрицей скоростей:

$$\begin{bmatrix} \mathcal{G}_1 \\ \mathcal{G}_2 \\ \mathcal{G}_3 \end{bmatrix} = \begin{bmatrix} -\sin(\varphi) & \cos(\varphi) & R \\ -\sin(\varphi + \Theta_2) & \cos(\varphi + \Theta_2) & R \\ -\sin(\varphi + \Theta_3) & \cos(\varphi + \Theta_3) & R \end{bmatrix} \cdot \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{bmatrix}. \quad (9)$$

Внутренняя инверсная кинематическая модель из выражения (9) может быть представлена в компактной форме  $\mathcal{G} = J \cdot \dot{q}$ . Иногда управление роботом происходит по его локальным координатам, которые можно получить, рассмотрев преобразование вращения исходя из соотношения:

$$\mathcal{G} = J \cdot (R_G^L)^T \cdot \dot{q}_M. \quad (10)$$

Прямая кинематическая модель в неподвижных координатах выражается из матрицы обратной кинематики (9) с учетом соотношения  $\dot{q} = S \cdot \mathcal{G}$ , где  $S = J^{-1}$ .

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{bmatrix} = \begin{bmatrix} -\sin(\Theta_1) & -\sin(\Theta_1 + \Theta_2) & -\sin(\Theta_1 + \Theta_3) \\ \cos(\Theta_1) & \cos(\Theta_1 + \Theta_2) & \cos(\Theta_1 + \Theta_3) \\ \frac{1}{2 \cdot R} & \frac{1}{2 \cdot R} & \frac{1}{2 \cdot R} \end{bmatrix} \cdot \begin{bmatrix} \mathcal{G}_1 \\ \mathcal{G}_2 \\ \mathcal{G}_3 \end{bmatrix}. \quad (11)$$

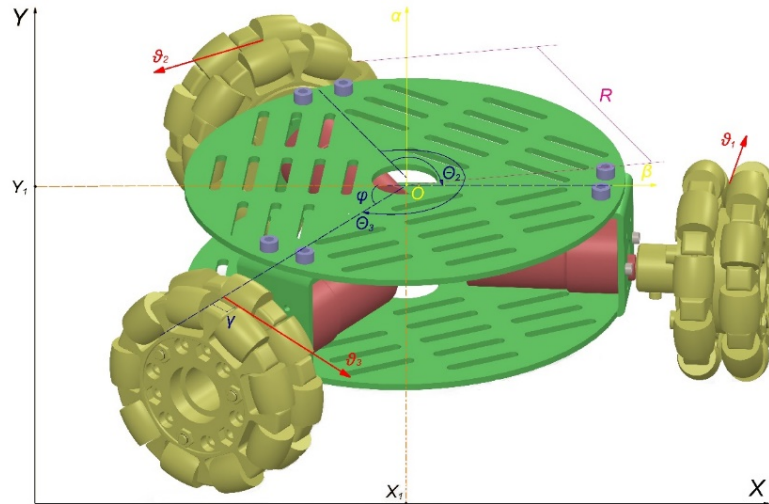


Рисунок 3. Пространственная кинематическая модель трехколесного КМР с *omni*-колесами

**Мобильный робот с колесами всенаправленного типа *mesanum*.**

Рассмотрим математическую модель колес всенаправленного типа в четырехколесных мобильных роботах. Одним из основных видов всенаправленных колес являются *mesanum*-колеса, представляющие собой колесо, на котором смонтировано несколько роликов, установленных на протяжении всей окружности обода. Угол вращения роликов составляет  $45^\circ$  между ними и осью колеса.

Кинематическая модель четырехколесного мобильного робота с *mesanum*-колесами изображена на рисунке 4.

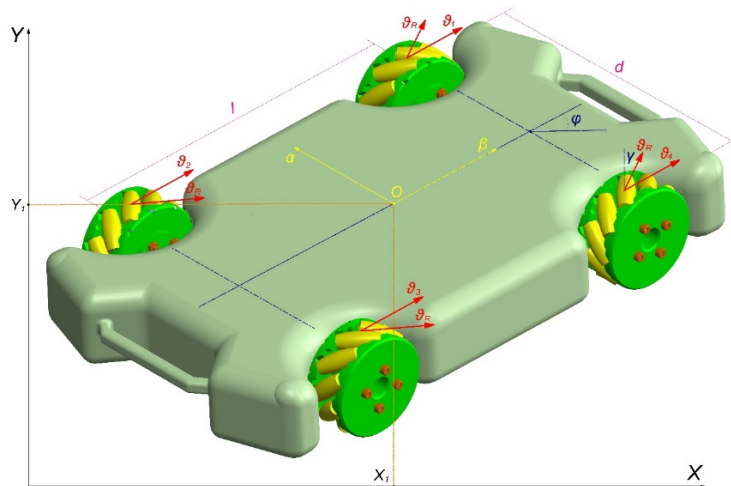


Рисунок 4. Кинематическая модель четырехколесного МРК с *mesanum*-колесами

Инверсная внешняя кинематическая модель четырехколесного мобильного робота с *mesanum*-колесами может быть разложена на несколько составляющих. Скорость переднего левого колеса в координатах плоскости  $(\vec{X}, \vec{Y})$  представляется суммой скоростей вращения главного колеса во времени  $\mathcal{G}_1(t)$  и скорости вращения роликов  $\mathcal{G}_R(t)$ .

Полная скорость движения робота в подвижной системе координат  $(\vec{\alpha}, \vec{\beta})$  определяется выражениями

$$\mathcal{G}_{\alpha 1} = \mathcal{G}_1 + \mathcal{G}_R \cdot \cos\left(\frac{\pi}{4}\right) = \mathcal{G}_1 + \frac{\mathcal{G}_R}{\sqrt{2}};$$

(12)

$$\mathcal{G}_{\beta 1} = \mathcal{G}_R \cdot \sin\left(\frac{\pi}{4}\right) = \frac{\mathcal{G}_R}{\sqrt{2}},$$

(13)

откуда находится скорость основного колеса:

$$\mathcal{G}_1 = \mathcal{G}_{\alpha 1} - \mathcal{G}_{\beta 1}.$$

(14)

Скорость переднего левого колеса в подвижной системе координат также может быть выражена через поступательную скорость робота:

$$\mathcal{G} = \sqrt{\dot{X}_1^2 + \dot{Y}_1^2}.$$

(15)

и его угловую скорость  $\dot{\varphi}$  посредством следующих выражений:

$$\mathcal{G}_{\alpha 1} = \dot{X}_1 - \dot{\varphi} \cdot \frac{d}{2};$$

(16)

$$\mathcal{G}_{\beta 1} = \dot{Y}_1 + \dot{\varphi} \cdot l,$$

(17)

где  $d$  – расстояние между колесами, расположенными на одной оси;  
 $l$  – длина платформы робота.

Из предыдущих уравнений следует, что скорость основного колеса может выражаться через скорость перемещения платформы робота:

$$\mathcal{G}_1 = \dot{X}_1 - \dot{Y}_1 - \left(l + \frac{d}{2}\right) \cdot \dot{\varphi}.$$

(18)

Аналогичные вычисления могут быть применены к остальным *tesanum*-колесам робота для определения их скоростей  $\mathcal{G}_2$ ,  $\mathcal{G}_3$ ,  $\mathcal{G}_4$ .

Инверсная кинематическая модель робота в неподвижной системе координат  $(\vec{X}, \vec{Y})$  определяется матричным методом

$$\begin{bmatrix} \mathcal{G}_1 \\ \mathcal{G}_2 \\ \mathcal{G}_3 \\ \mathcal{G}_4 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -\left(l + \frac{d}{2}\right) \\ 1 & 1 & -\left(l + \frac{d}{2}\right) \\ 1 & -1 & \left(l + \frac{d}{2}\right) \\ 1 & 1 & \left(l + \frac{d}{2}\right) \end{bmatrix} \cdot \begin{bmatrix} \dot{X}_1(t) \\ \dot{Y}_1(t) \\ \dot{\phi}(t) \end{bmatrix}. \quad (19)$$

Внутренняя инверсная кинематическая модель из уравнения (19) может быть представлена в компактной форме  $\mathcal{G} = J \cdot \dot{q}$ , где  $\mathcal{G}^T = [\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_4]^T$  и  $q^T = [X_1, Y_1, \phi]^T$ .

Для расчета инверсной кинематики в неподвижных координатах вводится матрица поворота, представляющая положение подвижной системы координат по отношению к подвижной ( $q = R_G^L q$ ),

$$R_G^L = \begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (20)$$

Тогда вектор скорости робота необходимо рассматривать как  $\mathcal{G} = JR_G^L \dot{q}$ .

Из выражения внутренней кинематической модели (9) в компактной форме, прямая внутренняя кинематическая модель представляется в виде  $q = J^+ \mathcal{G}$ , где  $J^+ = (J^T J)^{-1} J^T$  является псевдообратной матрицей  $J$ .

Прямая внутренняя кинематика четырехколесного мобильного робота с *тесанит*-колесами определяется матричным выражением:

$$\begin{bmatrix} \dot{X}_1 \\ \dot{Y}_1 \\ \dot{\phi} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ -\frac{1}{\left(l + \frac{d}{2}\right)} & -\frac{1}{\left(l + \frac{d}{2}\right)} & \frac{1}{\left(l + \frac{d}{2}\right)} & \frac{1}{\left(l + \frac{d}{2}\right)} \end{bmatrix} \cdot \begin{bmatrix} \mathcal{G}_1 \\ \mathcal{G}_2 \\ \mathcal{G}_3 \\ \mathcal{G}_4 \end{bmatrix}. \quad (21)$$

Прямая внутренняя кинематика в неподвижных координатах выражается в компактной форме уравнением  $\dot{q} = (R_G^L) J^+ v$ .

### Заключение.

Таким образом, применение колес всенаправленного типа в робототехнике повышает мобильность колесного робота за счет использования движения колеса в различных направлениях, что влечет увеличение числа математических моделей для описания управления

движения роботом, что повышает области применения данного типа колес в некоторых областях специального назначения.

### **Список литературы**

- [1] Кампион, Г. Структурные свойства и классификация кинематических и динамических моделей колесных мобильных роботов / Г. Кампион, Ж. Бастен, Б. д'Андреа-Новель // Нелинейная динамика. – 2011. – Т. 7, № 4. – С. 733–769.
- [2] Trojnacki, Maciej. Mechanical Properties of Modern Wheeled Mobile Robots / Maciej Trojnacki, Przemysław Dąbek // Journal of Automation, Mobile Robotics and Intelligent Systems. – 2019. – Vol. 13, № 3. – P. 3–13.
- [3] Klančar, Gregor. Wheeled Mobile Robotics. From Fundamentals Towards Autonomous Systems / Gregor Klančar, Andrej Zdešar, Sašo Blažič, Igor Škrjanc. – The Boulevard, Langford Lane, Kidlington, Oxford. – 2017. – 492 p.

## **CONTROL ALGORITHM FOR FOUR-WHEEL MOBILE ROBOTS**

**S.A. Pauliukavets**

*Head of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU, PhD of Technical Sciences, Associate Professor*

**A.A. Velchenko**

*Associate Professor of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU, PhD of Technical Sciences, Associate Professor*

**Wu Xinxin**

*PhD student of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU*

**A.A. Radkevich**

*Student of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU*

**N.O. Sauko**

*Student of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU*

*Department of «Electric drive and automation of industrial installations and technological complexes»  
Faculty of Information Technology and Robotics  
Belarusian National Technical University, Republic of Belarus  
E-mail: sap@bntu.by*

**Abstract.** In this paper, mathematical models for controlling robots with omnidirectional wheels (omni-wheel and mecanum-wheel) are considered, options for combining various types of omnidirectional wheels in a robot are presented.

**Keywords:** mobile robot, omnidirectional wheel, omni-wheel, mecanum-wheel, controllability, mobility.

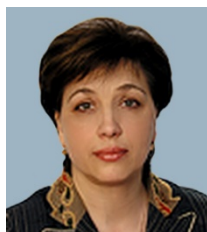


УДК 004.65

## КЛЮЧЕВЫЕ АСПЕКТЫ ВЫБОРА СУБД ДЛЯ АНАЛИТИКИ БОЛЬШИХ ДАННЫХ



**Н.В. Павлович**  
Магистрант кафедры  
проектирования  
информационно-  
компьютерных систем  
БГУИР  
traulovich00@gmail.com



**И.Н. Тонкович**  
Доцент кафедры  
проектирования  
информационно-  
компьютерных систем  
БГУИР, кандидат  
технических наук, доцент  
intonkovich@gmail.com



**В.Ф. Алексеев**  
Доцент кафедры  
проектирования  
информационно-  
компьютерных систем  
БГУИР, кандидат  
технических наук, доцент  
alexvikt.minsk@gmail.com

### **Н.В. Павлович**

Окончил БГУИР (2021 г.), в настоящее время является магистрантом этого университета. Проводит научные исследования в области распределенной обработки больших массивов данных.

### **И.Н. Тонкович**

Окончила Белорусский государственный университет, факультет прикладной математики. Основная область научных интересов связана с применением инновационных подходов в системе высшего образования, разработкой методов и алгоритмов построения информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

### **В.Ф. Алексеев**

Окончил Минский радиотехнический институт. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

**Аннотация.** В данной работе представлены ключевые аспекты выбора различных типов СУБД для аналитики больших данных. Рассмотрены критерии и факторы, влияющие на выбор СУБД. Выявлены особенности использования реляционных, нереляционных, графовых, колоночных и In-memory СУБД для аналитики Big Data.

**Ключевые слова:** аналитика, большие данные, СУБД для аналитики, реляционные, нереляционные, графовые, колоночные, In-memory.

### **Введение.**

Аналитика больших данных становится одним из ключевых направлений в современном бизнесе. По оценкам ResearchAndMarkets мировой рынок аналитики больших данных будет увеличиваться ежегодно в среднем на 11,9% с 2020 по 2028 годы. Такой рост проектов по внедрению систем бизнес-аналитики Big Data приведет к серьезному спросу на специализированные СУБД для обработки и анализа больших объемов данных. Выбор СУБД для аналитики Big Data является критически важной задачей, для решения которой необходимо поддерживать быстрые итерации для правильного анализа данных, учитывать способность обрабатывать структурированные и неструктурированные наборы данных, возможность масштабирования, поддержки параллельной обработки и распределенных вычислений, возможность работы с различными форматами данных, а также способность обеспечивать безопасность и защиту данных.

В данном исследовании рассмотрены ключевые аспекты выбора СУБД для аналитики Big Data.

### **Критерии и факторы, влияющие на выбор СУБД для аналитики Big Data.**

Выделим ряд общих критериев, которые следует учитывать при выборе СУБД для аналитики Big Data [1]:

- 1 Размер данных. СУБД должна быть способна обрабатывать большие объемы данных, а также поддерживать их хранение.
- 2 Скорость обработки данных. СУБД должна быть способна обрабатывать данные быстро и эффективно.
- 3 Сложность запросов. В зависимости от сложности запросов и аналитических операций необходимо выбирать СУБД с соответствующими функциональными возможностями.
- 4 Надежность и доступность. СУБД должна обеспечивать высокую надежность и доступность, чтобы обеспечить постоянный доступ к данным.
- 5 Стоимость внедрения и использования. Стоимость СУБД также является важным фактором при выборе, поскольку одни СУБД могут быть дороже, чем другие.
- 6 Совместимость с существующими инструментами. СУБД должна быть совместима с существующими инструментами аналитики и инфраструктурой.
- 7 Поддержка и развитие. СУБД должна иметь поддержку и обновления, а также продолжать развиваться и адаптироваться к новым технологиям.

Кроме того, при выборе СУБД для аналитики Big Data также важно учитывать ее тип (реляционная, NoSQL, NewSQL и т.д.), способность к параллельной обработке, возможность масштабирования и другие факторы.

Существуют различные типы СУБД, которые могут быть использованы для анализа Big Data. Каждый тип СУБД имеет свои особенности и применяется в зависимости от характеристик проекта и требуемых функциональных возможностей, типа данных и характера анализа. Более того, аналитика больших данных может быть проблематичной, поскольку она часто включает сбор и хранение различных типов неструктурированных данных или используется для обработки непрерывного потока данных в режиме реального времени.

Выделяют следующие типы СУБД для аналитики Big Data:

- 1 Реляционные СУБД (RDBMS) – это классические СУБД, которые используются для хранения и управления структурированными данными. Они могут использоваться для анализа больших данных, но часто не обладают достаточной масштабируемостью для обработки Big Data в режиме реального времени.
- 2 NoSQL СУБД – используются для хранения и обработки неструктурированных данных, таких как данные социальных сетей, IoT-данные и т.д. NoSQL СУБД более гибкие и масштабируемые, чем реляционные СУБД, что делает их популярным выбором для аналитики Big Data.
- 3 Графовые СУБД – как правило, используются для анализа связей между объектами в данных, таких как социальные сети, транспортные системы и т.д. Графовые СУБД могут быстро находить связи между данными, что делает их очень полезными для аналитики Big Data.
- 4 Колоночные СУБД – используются для хранения и обработки больших объемов структурированных данных. Колоночные СУБД могут обрабатывать большие объемы данных очень быстро и обладают высокой степенью сжатия данных, что делает их эффективными инструментами для аналитики Big Data.
- 5 In-memory СУБД – они используются для обработки данных в оперативной памяти компьютера, что обеспечивает высокую скорость обработки данных. In-memory СУБД обычно используются для анализа больших объемов данных в режиме реального времени.

Каждый тип СУБД имеет свои преимущества и недостатки, и выбор конкретного типа СУБД для аналитики Big Data зависит от требований проекта и доступных ресурсов.

### **Особенности использования реляционных СУБД для аналитики Big Data.**

Реляционные СУБД, такие как MySQL, Oracle и Microsoft SQL Server, долгое время были основным инструментом для работы с данными в больших компаниях. Они хорошо подходят

для обработки структурированных данных и отлично работают с небольшими и средними объемами информации.

Однако, с появлением Big Data, объемы и сложность обрабатываемых данных значительно возросли. Реляционные СУБД не всегда эффективно работают с такими объемами информации, и могут стать узким местом при анализе данных.

Тем не менее, реляционные СУБД все еще могут использоваться для аналитики Big Data, особенно в тех случаях, когда данные имеют ясную структуру и формат.

Для увеличения производительности и эффективности работы реляционных СУБД с Big Data, используются различные инструменты, такие как партиционирование, шардирование, репликация данных и т.д. Также в реляционных СУБД могут быть добавлены функции аналитики, такие как OLAP и Data Warehousing.

Однако, в целом, для работы с Big Data лучше использовать специализированные СУБД, которые оптимизированы для работы с большими объемами информации и могут обрабатывать неструктурированные данные.

С точки зрения практического внедрения или использования выделяют следующие наиболее популярные реляционные СУБД для аналитики Big Data:

1 Oracle Database – это одна из самых популярных реляционных СУБД в мире. Она обладает широким спектром возможностей для анализа данных и имеет хорошую масштабируемость для обработки больших объемов данных.

2 Microsoft SQL Server – это реляционная СУБД, разработанная Microsoft. Она имеет множество функций и возможностей для анализа данных, включая встроенный анализ данных и машинное обучение.

3 IBM Db2 – это еще одна популярная реляционная СУБД, которая используется для анализа Big Data. Она предлагает высокую производительность и масштабируемость, а также широкий набор функций и инструментов для анализа данных.

4 PostgreSQL – это бесплатная и открытая реляционная СУБД, которая может быть использована для анализа Big Data. Она имеет множество расширений и плагинов, которые могут быть использованы для расширения ее функциональности.

5 MySQL – это популярная бесплатная реляционная СУБД. Хотя она может иметь некоторые ограничения для обработки больших объемов данных, ее можно использовать для анализа небольших или средних объемов данных.

Это лишь некоторые из множества реляционных СУБД, которые могут быть использованы для аналитики Big Data. Выбор конкретной СУБД будет зависеть от конкретных потребностей и требований организации.

### **Особенности использование нереляционных СУБД для аналитики Big Data.**

Нереляционные СУБД, также известные как NoSQL базы данных, стали популярны в последние годы благодаря их способности обрабатывать большие объемы неструктурированных данных с высокой скоростью и масштабируемостью. Они используют различные модели данных и алгоритмы хранения, которые отличаются от традиционных реляционных баз данных.

Особенности использования нереляционных СУБД для аналитики Big Data:

1 Высокая масштабируемость. Нереляционные СУБД могут обрабатывать большие объемы данных и масштабироваться на несколько серверов, обеспечивая горизонтальное масштабирование.

2 Гибкость в модели данных. Нереляционные СУБД не требуют определения жесткой схемы базы данных, что облегчает добавление и изменение данных.

3 Более быстрый доступ к данным. Нереляционные СУБД могут обеспечить быстрый доступ к данным при использовании определенных алгоритмов хранения, таких как хранение в оперативной памяти или индексирование.

4 Поддержка неструктурированных данных. Нереляционные СУБД могут обрабатывать данные различных форматов, таких как документы, графы и временные ряды.

Нереляционные СУБД делятся на несколько категорий в зависимости от используемой модели данных, таких как ключ-значение, документоориентированные, столбцовые, графовые и др. Каждая категория подходит для конкретного типа данных и типичных запросов в аналитике Big Data.

Наиболее известными представителями нереляционных СУБД для аналитики Big Data являются [2]:

1 MongoDB – документоориентированная СУБД, которая хранит данные в формате JSON-подобных документов. Она хорошо подходит для работы с большими объемами неструктурированных данных.

2 Apache HBase – СУБД, работающая на основе Hadoop и HDFS. HBase подходит для хранения больших объемов данных с высокой производительностью в режиме реального времени.

3 Apache Couchbase – распределенная NoSQL СУБД, которая поддерживает хранение данных в оперативной памяти и на диске. Она может обрабатывать большие объемы структурированных и неструктурированных данных в режиме реального времени.

4 Amazon DynamoDB – управляемая NoSQL СУБД, которая хранит данные в виде ключ-значение. Она обеспечивает высокую доступность, производительность и масштабируемость при работе с большими объемами данных.

#### **Использование графовых и колоночных СУБД для аналитики Big Data.**

Графовые СУБД – это тип нереляционных СУБД, которые используют графовую модель данных для хранения и обработки информации. Графы представляют собой набор вершин (узлов), соединенных ребрами (связями), которые могут иметь различные свойства и атрибуты.

В аналитике Big Data графовые СУБД используются для анализа связей и зависимостей между данными, таких как социальные сети, транспортные сети, сети связи и т.д.

Одной из наиболее известных графовых СУБД для аналитики Big Data является Neo4j. Она представляет собой гибридную СУБД, которая сочетает в себе возможности графовой модели данных и языка запросов Cypher с традиционными реляционными возможностями, такими как ACID-транзакции и индексирование.

Графовые СУБД обычно имеют высокую производительность и масштабируемость для анализа больших объемов данных. Однако, они могут быть менее удобными для анализа данных, которые не являются связанными или зависимыми.

Колоночные СУБД являются специальным типом нереляционных СУБД, предназначенных для работы с большими объемами данных. Они используются для хранения и анализа структурированных данных, которые обычно представлены в виде больших таблиц. В колоночных СУБД данные хранятся в виде колонок, в отличие от реляционных СУБД, где они хранятся в виде строк.

Колоночные СУБД используются для аналитики Big Data, так как они обеспечивают быстрый доступ к большим объемам данных, позволяют проводить анализы в режиме реального времени и поддерживают параллельную обработку данных. Кроме того, они обладают высокой степенью сжатия данных, что позволяет существенно снизить объем хранимых данных.

Наиболее популярными решениями колоночных СУБД для аналитики Big Data являются:

1 Vertica – колоночная СУБД, разработанная для аналитической обработки больших объемов данных. Она поддерживает распределенную архитектуру и реализует функциональность параллельной обработки запросов, что позволяет обеспечить быстродействие и масштабируемость.

2 Amazon Redshift – облачная колоночная СУБД, разработанная для аналитической обработки больших объемов данных. Она использует распределенную архитектуру и высокопроизводительное хранилище данных, что позволяет обеспечить высокую скорость выполнения запросов и масштабируемость.

3 Google BigQuery – облачная колоночная СУБД, которая предоставляет масштабируемое хранилище данных и высокопроизводительный механизм выполнения запросов. Она использует технологии облачных вычислений, что позволяет обеспечить высокую доступность и масштабируемость.

4 Apache Cassandra – распределенная колоночная СУБД, разработанная для обработки больших объемов данных с высокой доступностью и масштабируемостью. Она использует распределенную архитектуру и реализует механизм репликации данных, что обеспечивает высокую отказоустойчивость.

5 ClickHouse – колоночная СУБД, разработанная для обработки больших объемов данных с высокой производительностью. Она использует распределенную архитектуру и механизмы параллельной обработки запросов, что позволяет обеспечить высокую скорость выполнения запросов и масштабируемость.

#### **Использование In-memory СУБД для аналитики Big Data.**

In-memory СУБД (IMDB) представляют собой базы данных, в которых данные хранятся в оперативной памяти компьютера. Такой подход позволяет быстро обрабатывать большие объемы данных, так как операции чтения и записи осуществляются непосредственно в памяти компьютера, а не на диске. Использование IMDB для аналитики Big Data позволяет существенно ускорить обработку данных и повысить скорость принятия решений.

IMDB часто используются в комбинации с другими СУБД, например, с колоночными или нереляционными СУБД, для обработки больших объемов данных. Кроме того, некоторые производители СУБД предлагают гибридные решения, которые объединяют в себе функциональность различных типов СУБД, включая IMDB.

Использование IMDB для аналитики Big Data может быть особенно полезным в случаях, когда требуется быстрый доступ к большим объемам данных, например, для аналитики рынка, финансового анализа или обработки больших объемов транзакций [3].

Топовые In-memory СУБД, которые можно использовать для аналитики Big Data:

1 SAP HANA – высокопроизводительная In-memory СУБД, которая может использоваться для анализа больших объемов данных в режиме реального времени.

2 Oracle TimesTen – In-memory СУБД, предназначенная для обработки транзакций и анализа данных в режиме реального времени.

3 IBM SolidDB – In-memory СУБД с открытым исходным кодом, которая обеспечивает быстрый доступ к данным и возможность анализа больших объемов данных.

4 MemSQL – In-memory СУБД с поддержкой SQL, которая может использоваться для анализа больших объемов данных в режиме реального времени.

5 Apache Ignite – распределенная In-memory СУБД, которая может использоваться для анализа больших объемов данных, распределенных по нескольким узлам.

6 Aerospike – In-memory СУБД, которая может использоваться для анализа больших объемов данных в режиме реального времени и обеспечивает высокую доступность и масштабируемость.

7 Altibase – In-memory СУБД, которая может использоваться для анализа больших объемов данных в режиме реального времени и предоставляет поддержку SQL и NoSQL.

#### **Заключение.**

В заключении можно отметить, что выбор СУБД для аналитики Big Data является важным и ответственным шагом. Необходимо учитывать ряд критериев, таких как тип данных, производительность, масштабируемость и прочее. Для аналитики Big Data могут использоваться различные типы СУБД, такие как реляционные, нереляционные, гибридные, графовые, колоночные и In-memory СУБД. Каждый тип имеет свои преимущества и недостатки, и выбор должен основываться на конкретных потребностях бизнеса.

### Список литературы

[1] Выбор базы данных для аналитики [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/487498/>.

[2] NoSQL [Электронный ресурс]. – Режим доступа: <https://www.bigdataschool.ru/wiki/nosql>.

[3] Использование In-memory в Big Data [Электронный ресурс]. – Режим доступа: <https://www.bigdataschool.ru/blog/tarantool-use-cases-and-advantages.html>.

## KEY ASPECTS OF CHOOSING A DBMS FOR ANALYTICS BIG DATA

***N.V. Paulovich***

*Master student of the  
Department of Information  
and Computer Systems  
Design BSUIR*

***I.N. Tonkavich***

*Associate Professor, Department  
of Information Computer Systems  
Design, PhD of Chehnical  
sciences, Associate Professor*

***V.F. Alekseev***

*Associate Professor, Department  
of Information Computer Systems  
Design, PhD of Technical sciences,  
Associate Professor*

*Department of Information and Computer Systems Design*

*Faculty of Computer Engineering*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: mpaulovich00@gmail.com, intonkovich@gmail.com, alexvikt.minsk@gmail.com*

**Abstract.** This paper presents the key aspects of choosing different types of DBMS for big data analytics. Criteria and factors influencing the choice of DBMS are considered.

The features of using relational, non-relational, graph, column and In-memory DBMS for Big Data analytics are revealed.

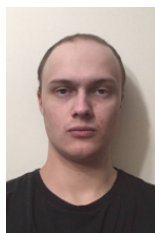
**Keywords:** analytics, big data, DBMS for analytics, relational, non-relational, graph, column, In-memory.

УДК 004.042

## DATA ANALYTICS В ЧАТАХ



**Ю.П. Плахотникова**  
студент БГУИР  
pocketickchaya@gmail.com



**Т.С. Ширяев**  
студент БГУИР  
tim.shiryaev@gmail.com



**С.Н. Нестеренков**  
доцент  
кафедры программного обеспечения  
информационных технологий, декан  
факультета компьютерных систем  
и сетей, кандидат технических наук  
s.nesterenkov@bsuir.by

### **Ю.П. Плахотникова**

Студент 4 курса специальности “Информатика и технологии программирования” БГУИР.

### **Т.С. Ширяев**

Студент 4 курса специальности “Информатика и технологии программирования” БГУИР.

### **С.Н. Нестеренков**

Кандидат технических наук, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры Программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

**Аннотация.** Рассмотрены методы анализа данных в чатах, которые используются для изучения и оценки продукта в целях его совершенствования. Описывается процесс сбора данных, их обработки и представления. Описаны принципы и технологии машинного обучения, которые можно применять для анализа данных в чатах и их преимущества. С помощью примеров из реальной жизни показывается, как анализ данных в чатах может быть полезен для бизнеса.

**Ключевые слова:** анализ данных, машинное обучение, чаты.

### **Введение.**

Анализ данных – это процесс изучения, очистки, преобразования и моделирования данных с целью получения полезной информации и поддержки принятия решений. Этот процесс включает в себя использование математических методов, статистических алгоритмов и компьютерных технологий. Анализ данных — это ключевой элемент многих областей деятельности, например, в бизнесе, медицине, науке и общественных науках.

В бизнесе анализ данных часто используется для прогнозирования спроса на товары и услуги, определения цен и маркетинговых стратегий. Для анализа данных в бизнесе может использоваться множество инструментов, включая макросы Excel, программное обеспечение для бизнес-аналитики, базы данных и инструменты машинного обучения.

Современными способами общения являются чаты в мессенджерах, социальных сетях и других приложениях. Анализ данных в чатах играет огромную роль в современных коммуникациях. Это связано с тем, что каждый день миллионы людей по всему миру используют чаты для онлайн-общения, работы, обучения и даже для прохождения медицинских консультаций. Многие из этих сообщений содержат важную информацию, которая может быть использована для анализа, определения различных показателей и получения ценной информации. Анализ данных в чатах — это процесс извлечения информации из текстовых сообщений, которые отправляют пользователи. Это может



включать в себя анализ тональности, определение ключевых слов и тематики, анализ эмоций и т.д. Анализ тональности позволяет определить, каким образом пользователи выражают свои эмоции и отношения к определенным событиям. Например, если пользователи обсуждают новинку, то можно узнать, каковы их отзывы - положительные, отрицательные или нейтральные. Анализ тональности может дать представление общего отношения к бренду, товару или услуге, что может использоваться для принятия решений по улучшению продукта или коммуникации с клиентами.

Ключевые слова и тематика также важны при анализе данных в чатах. Количество и частота использования определенных слов может дать представление об интересах и потребностях пользователей. Например, если пользователи часто упоминают слова "путешествия", "отпуск" и "туризм", то можно сделать вывод, что у них есть интерес к этим темам. Данные могут быть использованы для персонализации сообщений и предложений, чтобы лучше соответствовать потребностям пользователей. Анализ эмоций позволяет выявить какие эмоции вызывает определенный продукт, событие или бренд. Эта информация может быть использована, чтобы сделать более эффективные маркетинговые кампании и улучшить общий имидж компании.

Важно понимать, что анализ данных в чатах может быть сложным и требует использования специальных инструментов и технологий. Однако, если правильно провести анализ и извлечь ценную информацию, то это может помочь улучшить качество продукта, снизить затраты и усилить взаимодействия с клиентами.

В целом, анализ данных в чатах — это мощный способ получения ценной информации о потребностях и интересах пользователей. Это может помочь компаниям улучшить свою продукцию и услуги, а также оптимизировать маркетинговые кампании.[1]

### **Оперативный анализ.**

По мере того, как компании начинают осознавать ценность анализа текстового контента, содержащегося в чате, они делают это двумя разными способами:

1. Анализ пользовательского чата. Эта методология анализа фокусируется на сообщениях, отправленных клиентом, чтобы раскрыть темы, возникающие проблемы, обсуждаемые клиентской базой.

2. Анализ чата пользователя с технической поддержкой. Этот метод анализа фокусируется на мониторинге поведения агентов поддержки, чтобы обеспечить соблюдение протоколов и определить области для улучшения ПО.

Разделив анализ на эти два разных способа, компании получают больше возможностей для улучшения деловой практики как внутри компании, так и за ее пределами. Также с точки зрения реализации анализ данных в чатах можно разделить на несколько категорий:

1. Использование алгоритмов машинного обучения. Алгоритмы машинного обучения способны проанализировать большой объем данных с высокой точностью. Они могут идентифицировать изображения, тексты и другой контент, чтобы получить полезную информацию. Модели машинного обучения умеют распознавать формат и структуру сообщений, а также делать выводы на основе анализа текста.

2. Анализ данных в режиме реального времени. Анализ данных в режиме реального времени позволяет мгновенно получать информацию о чатах по мере их появления. Режим реального времени актуален, когда нужно реагировать на данные мгновенно. Этот способ анализа данных может повысить эффективность работы чатов. Системы, работающие в режиме реального времени, могут быстро и эффективно обрабатывать даже огромные объемы данных и выдавать результаты анализа в мгновение ока. Исключительно важно, чтобы данные поступали в систему быстро, иначе система не сможет реагировать корректно.

3. Анализ, включающий в себя работу эмоционального интеллекта. Эмоциональный интеллект — это способность человека понимать свои и чужие эмоции, управлять ими и использовать в процессе общения. Аналитика на основе эмоционального интеллекта может быть использована для идентификации сильных и слабых сторон в коммуникации пользователей в чате. Она может обратить внимание на нарушения, убыль эмоциональной интенсивности, перебои в общении и необходимость корректировки манеры общения, что может помочь в разработке ИИ.

#### **Центральные темы.**

Лучшее место для начала анализа текста данных чата — это определение основных тем с помощью комбинаций ключевых слов и категорий. Ключевые слова — это парные сочетания слов, определяемые по частоте их употребления. Форматы облака тегов обеспечивают простую визуализацию данных для определения наиболее часто упоминаемых терминов в контенте. С помощью облаков тегов легко определить, какие типы продуктов или названия (например, Money Market или Debit Card) часто обсуждаются. Категории (типы) разговоров автоматически определяются платформой на основе множества факторов, включая частоту, силу и качество словесных ассоциаций в данных чата, настроение и другие. Они отображаются в порядке статистической значимости. Визуализация как категорий, так и комбинаций ключевых слов позволяет отобразить наиболее важные области, требующие внимания.

#### **Тенденции.**

Следующим шагом в анализе данных чата является поиск тенденций в каждой категории и в наиболее часто используемых словосочетаниях. Важно знать, повышается или понижается объем и общее впечатление по каждой теме месяц за месяцем. Например, финансовый отдел может столкнуться с ошибками, возникающими при использовании продукта, и ростом ряда вопросов, связанных с налогообложением, весной, достигающими пика в апреле. Тенденции текстовой аналитики также могут определять вопросы, возникающие при использовании продукта, и комментарии о его обновлениях. Возможно, была запущена новая версия мобильного приложения, что вызвало увеличение количества обсуждений программного обеспечения. Предприятия могут внедрять планы совершенствования процессов и ожидать от клиентов подтверждения того, что их усилия принесут желаемые результаты. В любом случае тенденции данных в диалогах покажут прогресс в разработке продукта.

#### **Постоянные проблемы.**

Текстовая аналитика чата также может выявить проблемы, которые регулярно возникают у клиентов. Примеры включают часто задаваемые вопросы, которые не рассматриваются на странице часто задаваемых вопросов веб-сайта. Могут быть трудности с определенными версиями браузера или операционной системы, вызывающие разное взаимодействие с пользователем. Стандартный процесс бизнес-транзакций может быть источником поддержки в чате. Изучая данные сеансов чата, можно выявить эти постоянные проблемы и глубже изучить переменные, такие как географическое распределение. Изучение отдельных сеансов чата поможет подтвердить догадки и достичь лучших результатов. Также важно искать «социальных лидеров мнений» в данных диалогов и убедиться, что их проблемы решены на 100% [2].

В качестве улучшения программного обеспечения на основе полученных в ходе анализа данных можно обеспечить:

- Дополнительные функции/возможности в продукте
- Обновление часто задаваемых вопросов на сайте
- Улучшение пользовательского интерфейса
- Улучшение обмена сообщениями об ошибках и обработки ошибок
- Корректировка взаимодействия сотрудников с клиентской базой

— Внедрение перспективного плана анализа данных для выявления проблем практически в реальном времени.

**Заключение.**

Анализ набора неструктурированных данных сеансов чата может напрямую привести к повышению эффективности бизнеса. Текстовая аналитика не только выявит конкретные проблемные области, но и подтвердит, а также обеспечит успешные обновления ПО.

**Список литературы**

[1] Вайгенд Андреас. BIG DATA. Вся технология в одной книге / Андреас Вайгенд. – Москва: Эксмо, 2018.–384 с.

[2] Радченко, И.А. Технологии и инфраструктура Big Data / И.А. Радченко, И.Н. Николаев. – СПб: Университет ИТМО, 2018.–52 с.

**DATA ANALYTICS IN CHATS**

***Y.P. Plakhotnikova***  
*Student of BSUIR*

***T.S. Shiryayev***  
*Student of BSUIR*

***S.N. Nesterenkov***  
*Candidate of Technical Sciences,*  
*Associate Professor of the Department*  
*of Information Technology Software,*  
*Dean of the Faculty of Computer*  
*Systems and Networks*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*  
*E-mail: packetickchaya@gmail.com, tim.shiryayev@gmail.com, s.nesterenkov@bsuir.by*

**Abstract.** The methods of data analysis in chats, which are used to study and evaluate the product in order to improve it, are considered. The process of data collection, processing and presentation is described. The principles and technologies of machine learning that can be used to analyze data in chats and their advantages are described. Using real-life examples, it shows how data analysis in chats can be useful for business.

**Keywords:** data analysis, machine learning, chats.

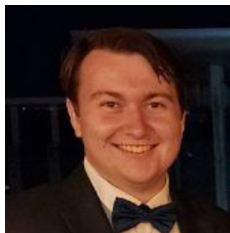
УДК 004.622

## АНАЛИЗ СОЦИАЛЬНЫХ СЕТЕЙ С ПОМОЩЬЮ BIG DATA: ВЫЯВЛЕНИЕ ТРЕНДОВ



**М.В. Реут**

Студент 5 курса, кафедры информатики, инженер-системный программист  
[maksreut1@gmail.com](mailto:maksreut1@gmail.com)



**В.Д. Владыцев**

Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР  
[v.vladymtsev@bsuir.by](mailto:v.vladymtsev@bsuir.by)



**А.Н. Марков**

Старший преподаватель, заместитель начальника Центра информатизации и инновационных разработок БГУИР  
[a.n.markov@bsuir.by](mailto:a.n.markov@bsuir.by)

### **М.В. Реут**

Студент 5 курса “Информатика и технологии программирования” БГУИР.

### **В.Д. Владыцев**

Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР

### **А.Н. Марков**

Старший преподаватель кафедры ПОИТ, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Использование методов big data для анализа социальных сетей с целью выявления трендов и мнений пользователей. Описывается процесс сбора, обработки и анализа данных, а также основные инструменты и методы, применяемые при работе с большими данными. Рассматривается пример анализа социальной сети Signal и дается оценка этических и правовых аспектов использования таких методов. В заключении подводятся итоги и указывается на значимость анализа социальных сетей в различных сферах деятельности.

**Ключевые слова:** big data, анализ социальных сетей, выявление трендов, сбор данных, обработка данных, инструменты и методы, этические и правовые аспекты, Signal, социальные сети, данные, машинное обучение, алгоритмы, статистический анализ, текстовый анализ, сентимент-анализ, визуализация данных, рекомендательные системы, пользователи, технологии, маркетинг, медиа.

### **Введение.**

Социальные сети играют все более важную роль в нашей повседневной жизни, и с каждым днем становится все сложнее обработать огромное количество информации, создаваемой ежедневно миллионами пользователей. В этом контексте анализ социальных сетей с помощью больших данных становится все более важным, чтобы выявлять тренды и мнения, которые могут помочь нам понять, что происходит в нашем обществе. В данном исследовании мы рассмотрим анализ социальной сети Signal с помощью big data. Представим план, который поможет провести анализ и выявить наиболее популярные темы, настроения пользователей и тренды в использовании данной социальной сети. Начнем со сбора данных из Signal с помощью API, который предоставляет данная социальная сеть, затем обработаем данные и проведем анализ, используя методы машинного обучения и анализа данных. Наконец, мы проанализируем результаты и сделаем выводы на основе данных, полученных из Signal. Анализ социальных сетей с помощью big data может помочь в понимании, как пользователи используют эту социальную сеть, что в свою очередь может помочь в улучшении ее функциональности и предоставлении более полезного опыта для пользователей.

### **Сбор данных в социальных сетях.**

Сбор данных в социальных сетях - это первый и один из самых важных этапов анализа больших данных. В случае с Signal, сбор данных может быть выполнен с помощью API, предоставляемого этой социальной сетью. Первый шаг заключается в определении целевых групп и тем, которые мы хотим проанализировать в Signal. Это может быть сделано, например, путем определения определенных ключевых слов, хештегов или пользователей, которые являются наиболее интересными для исследования. Для того чтобы получить доступ к API необходимо зарегистрироваться на сайте разработчика Signal, получить ключ API и настроить нашу систему для взаимодействия с Signal. После получения доступа к API можно начать сбор данных. Для этого нужно определить параметры, которые нужно получить, например, текст сообщений, даты публикации, количество лайков и репостов и т.д. После сбора данных необходимо их обработать, чтобы избавиться от ненужных данных, провести их очистку и предварительный анализ. Для этого можно использовать специализированные программные инструменты, такие как Python, R и другие. После обработки данных необходимо их сохранить в базу данных или файловую систему, чтобы иметь возможность их использовать в будущем для анализа.

Что касается других мессенджеров и социальных сетей, то Twitter и Facebook продолжают предоставлять доступ к своим API разработчикам, но в свете недавних скандалов с утечкой персональных данных и давлением со стороны правительств на защиту конфиденциальности, они стали более осторожными и регулируют доступ к своим данным. Однако в связи с последними новостями Twitter планирует запустить платный доступ к своему API и стоимость базового доступа к программному обеспечению составит 100\$ в месяц. WhatsApp, с другой стороны, ограничила доступ к своему API в 2021 году, чтобы защитить конфиденциальность своих пользователей и предотвратить распространение мошенничества и спама.

### **Обработка данных.**

После сбора данных необходимо провести их очистку, чтобы удалить шумы и ненужную информацию, такую как рекламные сообщения, боты и т.д.

### **Анализ данных.**

Анализ данных социальных сетей включает в себя несколько шагов. На первом этапе необходимо определить насколько популярна тема, которую исследуем, и выявляем тренды в качестве упоминаний. Затем необходимо анализировать пользователей, выделяя группы с общими интересами и оценивая их влияние. На третьем этапе необходимо анализировать содержимое, выделяя наиболее часто используемые слова и фразы, а также оценивая тональность текстов. После этого используется визуализация данных для наглядного представления результатов анализа.

### **Сравнительный анализ Signal с другими мессенджерами.**

Каждый мессенджер и социальная сеть имеют различные политики конфиденциальности и подходы к обработке и использованию пользовательских данных в научных исследованиях. Некоторые из них предоставляют открытые API (Application Programming Interface), которые разрешают ученым получать доступ к определенным данным, в то время как другие ограничивают доступ к данным и обрабатывают их сами. Например, Twitter предоставляет доступ к своим данным через открытое API, что позволяет исследователям получать данные о твитах, пользовательских профилях и тенденциях. Twitter также получает прибыль от продажи данных пользователей, в том числе информации об их интересах и поведении в социальной сети.

Компания предоставляет доступ к этим данным рекламодателям, которые используют их для создания более эффективных рекламных кампаний. В то же время, Twitter также заявляет, что защищает конфиденциальность данных пользователей и позволяет им контролировать свою конфиденциальность через настройки приватности в своих аккаунтах. Однако, подобные политики использования данных часто становятся объектом обсуждения с точки зрения этики и законности. Facebook также предоставляет открытый доступ к своим данным, но только после

прохождения процедуры аутентификации и получения соответствующих разрешений. Однако, Facebook была критикована за то, что она не защищает свои данные пользователя, в результате чего был создан крупный скандал в 2018 году. WhatsApp, в отличие от Facebook, не позволяет исследователям получить доступ к данным своих пользователей через открытые API и, как следствие, ученые обычно используют альтернативные методы для сбора данных. Signal, с другой стороны, считается одним из наиболее безопасных мессенджеров, так как вся коммуникация между пользователями зашифрована, и компания не продает данные своих пользователей.

Features	DMME	Whatsapp	Telegram	Viber	WeChat	Signal
Military-grade encryption	✓	✗	✗	✗	✗	✓
Encrypted File Transfer	✓	✗	✗	✓	✗	✓
No phone number required	✓	✗	✗	✗	✗	✗
Self-Destructed Messages	✓	✗	✓	✓	✗	✗
Local stored data (Encrypted on your phone & not stored on servers)	✓	✗	✗	✗	✗	✗
Peer-to-peer payments	✓	✗	✓	✗	✓	✗
Public / Private Key (email=pubkey / psw=privkey)	✓	✗	✓	✓	✗	✓

Рисунок 1. Сравнение безопасности мессенджеров

### Выявление трендов и мнений.

Для выявления трендов и мнений в сообществе Signal можно использовать методы анализа тональности текстов, тематического моделирования и анализа социальной сети. Методы анализа тональности позволят определить, какие слова и выражения используются в сообщениях с положительной, отрицательной или нейтральной окраской. Тематическое моделирование поможет выделить ключевые темы и подтемы, обсуждаемые в сообществе Signal. Анализ социальной сети позволит определить наиболее активных пользователей и группы, а также выявить связи между ними.

### Визуализация результатов.

Для наглядной визуализации результатов анализа можно использовать различные инструменты, такие как графики, диаграммы и облака слов. Графики и диаграммы помогут показать изменение количество сообщений и активности пользователей во времени, а также распределение сообщений по тональности и темам. Облака слов позволят выделить наиболее употребляемые слова и выражения в сообщениях.

### Интерпретация результатов.

Полученные результаты анализа социальной сети Signal помогут выявить ключевые тренды и мнения, обсуждаемые в сообществе, а также определить наиболее активных пользователей и группы. Эта информация может быть полезна для маркетинговых

исследований, анализа общественного мнения и разработки стратегий коммуникации в социальных сетях.

#### **Заключение.**

В заключении можно отметить, что анализ социальных сетей с помощью Big Data является важным инструментом для выявления трендов и мнений. В случае социальной сети Signal, были рассмотрены процесс сбора, обработки и анализа данных, показав, что при правильном подходе и использовании соответствующих инструментов, можно получить ценные знания о поведении пользователей и их мнениях. Однако, необходимо учитывать, что сбор и анализ данных социальных сетей связаны с этическими и правовыми вопросами, такими как конфиденциальность и защита личных данных. Поэтому, при использовании таких методов необходимо соблюдать правила и законы, чтобы избежать негативных последствий.

Тем не менее анализ социальных сетей остается важным инструментом для многих областей, от маркетинга и бизнеса до научных исследований и политической аналитики. И при правильном использовании этого инструмента можно получить ценную информацию, которая поможет принимать более обоснованные решения и добиваться успеха в различных областях деятельности.

#### **Список литературы**

- [1] Вайгенд Андреас. Big Data: революция социальных данных.
- [2] Джозел Грас. Data Science: Наука о данных с нуля.
- [3] Ольга Базалева. Мастерство визуализации данных
- [4] Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662-679.
- [5] Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.

## **SOCIAL MEDIA ANALYSIS USING BIG DATA: IDENTIFYING TRENDS**

**M.V. Reut**

*Student of BSUIR, software engineer*

**V. D. Vladymtsev**

*Assistant of the Department of Computer Science, Software Engineer of DIACS CIIR BSUIR*

**A.N. Markov**

*Senior lecturer of the department, Deputy head of the Center for Informatization and Innovative Developments*

*Center for Informatization and Development of the Belarusian University of State Informatics and Radioelectronics, Republic of Belarus*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*

*E-mail: maksreut1@gmail.com, v.vladymtsev@bsuir.by, a.n.markov@bsuir.by*

**Abstract.** Using big data methods to analyze social networks in order to identify trends and user opinions. It describes the process of data collection, processing and analysis, as well as the main tools and methods used when working with big data. An example of Signal social network analysis is considered and an assessment of the ethical and legal aspects of using such methods is given. In conclusion, the results are summarized and the importance of the analysis of social networks in various fields of activity is indicated.

**Keywords:** big data, social network analysis, trend detection, data collection, data processing, tools and methods, ethical and legal aspects, Signal, Social networks, Data, machine learning, algorithms, statistical analysis, text analysis, sentiment analysis, data visualization, recommendation systems, users, technologies, marketing, media.



УДК 621.311.25-027.45

## ИНФОРМАЦИОННАЯ СРЕДА ДЛЯ ОПТИМИЗАЦИИ НАДЗОРА В ОБЛАСТИ ОБЕСПЕЧЕНИЯ ЯДЕРНОЙ И РАДИАЦИОННОЙ БЕЗОПАСНОСТИ ПРИ ЭКСПЛУАТАЦИИ БЕЛОРУССКОЙ АЭС



**С.М. Сацук**

Заведующий кафедрой электроники БГУИР,  
кандидат технических наук, доцент  
satsuk@bsuir.by



**С.В. Дробот**

Декан факультета информационной  
безопасности БГУИР, кандидат технических  
наук, доцент



**В.Н. Русакович**

Старший преподаватель кафедры электроники  
БГУИР



**Д.О. Серов**

Студент БГУИР

### **С.М. Сацук**

Окончил Минский радиотехнический институт. Область научных интересов: обоснование и экспертиза безопасности в области использования атомной энергии и источников ионизирующего излучения; технология формирования изделий микро-, нано- электроники с использованием вентильных металлов и их анодных оксидов, легированных редкоземельными металлами.

### **С.В. Дробот**

Окончил Минский радиотехнический институт. Область научных интересов: обоснование и экспертиза безопасности в области использования атомной энергии и источников ионизирующего излучения; дифракция электромагнитных волн; математическое моделирование и разработка устройств СВЧ-диапазона; автоматизированное проектирование СВЧ-устройств в интегральном исполнении; измерение электрофизических параметров материалов в СВЧ-диапазоне.

### **В.Н. Русакович**

Окончил Минский радиотехнический институт. Область научных интересов: обоснование и экспертиза безопасности в области использования атомной энергии и источников ионизирующего излучения; дифракция электромагнитных волн; измерение электрофизических параметров материалов в СВЧ-диапазоне; математическое моделирование устройств СВЧ-диапазона.

### **Д.О. Серов**

Студент 4 курса специальности «Электронные и информационно-управляющие системы физических установок» БГУИР.

**Аннотация.** Создана информационная среда для сбора, накопления, хранения и анализа показателей безопасной эксплуатации Белорусской АЭС.

Показано, что информационная среда позволит повысить эффективность надзорной (контрольной) деятельности в области обеспечения ядерной и радиационной безопасности в Республике Беларусь и внедрить передовые технологии в планирование и осуществление регулирующей деятельности на всех этапах надзора за Белорусской АЭС.

**Ключевые слова:** информационная среда, ядерная и радиационная безопасность, АЭС, нормативные правовые акты.

### **Введение.**

Белорусская АЭС является объектом, на площадке которого установлен особый порядок организации и осуществления надзора в области обеспечения ядерной и радиационной безопасности.

Такой порядок распространяется на все этапы жизненного цикла АЭС, установленные законодательством в области использования атомной энергии.

Безопасность АЭС достигается за счет качественного проектирования, конструирования и изготовления оборудования, размещения, сооружения и эксплуатации АЭС, соблюдения требований законодательства в области использования атомной энергии, формирования и поддержания культуры безопасности, учета опыта эксплуатации и современного уровня развития науки, техники и производства [1].

Для обеспечения выполнения указанных принципов необходимо использовать новые технологии и, в частности, такой мощный инструмент как «Большие данные» (Big Data).

Наличие технологий обработки данных больших объемов является ключевым фактором успеха для целого ряда применений, включая проектирование и эксплуатацию атомных электростанций, обращение с ядерными отходами, исследование термоядерного синтеза, мониторинг окружающей среды, контроль материалов и другие.

### **Актуальность.**

Применительно к надзору в области обеспечения ядерной и радиационной безопасности при эксплуатации Белорусской АЭС, предполагающему использование показателей эффективности эксплуатационной деятельности, важным элементом обеспечения ядерной и радиационной безопасности является процесс сбора, накопления, хранения, анализа и представления структурированных и неструктурированных данных по показателям безопасной эксплуатации.

Аналитика большого объема информации позволит автоматизировать работу сотрудников надзорного органа, Белорусской АЭС, повысить эффективность их работы, направленной на обеспечение ядерной и радиационной безопасности.

### **Надзорная деятельность в области обеспечения ядерной и радиационной безопасности Белорусской АЭС.**

В документе МАГАТЭ [2] определены основополагающая цель безопасности и десять соответствующих принципов безопасности.

Основополагающая цель безопасности – защита людей и охрана окружающей среды от вредного воздействия ионизирующих излучений.

Принципы безопасности применяются в соответствующих случаях на протяжении всего жизненного цикла ядерных установок, используемых в мирных целях, а также в отношении защитных мер для уменьшения существующих радиационных рисков.

Они обеспечивают основу для разработки требований и мер, предназначенных для обеспечения защиты людей и охраны окружающей среды от радиационных рисков, а также обеспечения безопасности ядерных установок и деятельности, связанных с этими рисками.

В частности, один из принципов безопасности, определенный в документе [2], гласит, что необходимо создавать и совершенствовать систему руководства и управления в интересах обеспечения безопасности в организациях, занимающихся радиационными рисками, и на установках, и в рамках деятельности, связанных с радиационными рисками.

Процесс оценки безопасности – это непрерывный, повторяющийся и длительный процесс, связанный с накоплением большого объема информации по опыту эксплуатации АЭС, ее анализом и выработкой мер, направленных на обеспечение безопасности.

Использование разрабатываемой информационной среды позволит реализовать основные требования документа [2] по обеспечению сотрудников контролирующих и регулирующего органов инструментарием для совершенствования надзорной деятельности.

Разрабатываемая информационная среда предназначена для выполнения следующих основных функций:

- сбор, накопление, хранение и анализ данных по показателям безопасной эксплуатации Белорусской АЭС на специальных сетевых файловых хранилищах;
- хранение данных в непротиворечивой и целостной форме в требуемых форматах;
- совместное использование данных многими пользователями;
- защита данных от несанкционированного доступа, искажения и/или уничтожения;
- разделение и предоставление прав доступа различным категориям пользователей;
- поиск информации и ее сортировку.

Для повышения надежности функционирования информационной среды предусмотрена реализация функций сохранения автономности сервера и клиента, отсутствие прямого доступа к содержимому информационной среды.

Также учтены ситуации, связанные со сбоем в электроснабжении в различных элементах информационной среды, в том числе и сбой программного обеспечения сервера.

Проектирование информационной среды осуществлялось с учетом требований нормативных правовых актов, в том числе технических нормативных правовых актов Республики Беларусь [3], международных [4-5], в том числе российских [6].

С учетом того, что информационная среда рассчитана на работу в течение всего срока эксплуатации Белорусской АЭС, который составляет 60 лет, то при проектировании большое внимание было уделено сокращению избыточности хранимых данных.

#### **Заключение.**

Создание информационной среды для сбора, накопления, хранения и анализа показателей безопасной эксплуатации Белорусской АЭС позволит повысить эффективность надзорной (контрольной) деятельности в области обеспечения ядерной и радиационной безопасности в Республике Беларусь и внедрить передовые технологии в планирование и осуществление регулирующей деятельности на всех этапах надзора за Белорусской АЭС.

#### **Список литературы**

[1] НП ЯРБ Общие положения обеспечения безопасности атомных электростанций, утвержденные постановлением МЧС Республики Беларусь от 13.04.2020 № 15. [Электронный ресурс]. – Режим доступа: [https://gosatomnadzor.mchs.gov.by/upload/iblock/d06/postmes\\_15\\_20200413\\_red\\_20200730.pdf](https://gosatomnadzor.mchs.gov.by/upload/iblock/d06/postmes_15_20200413_red_20200730.pdf) (дата обращения: 17.04.2023).

[2] основополагающие принципы безопасности. Нормы безопасности МАГАТЭ. Основы безопасности. № SF-1 // МАГАТЭ, Вена, 2007. [Электронный ресурс]. – Режим доступа: [https://www-pub.iaea.org/MTCD/Publications/PDF/pub1273r\\_web.pdf](https://www-pub.iaea.org/MTCD/Publications/PDF/pub1273r_web.pdf) (дата обращения: 17.04.2023).

[3] Постановление Министерства по чрезвычайным ситуациям Республики Беларусь от 06.09.2021 № 59 «О требованиях к содержанию отчета по оценке текущего состояния безопасности». [Электронный ресурс]. – Режим доступа: <https://pravo.by/document/?guid=12551&p0=W22137225&p1=1> (дата обращения: 17.04.2023).

[4] Государственная, правовая и регулирующая основа обеспечения безопасности. Нормы безопасности МАГАТЭ. Общие требования безопасности. № GSR Part 1 (Rev. 1). Вена: МАГАТЭ, 2016. – Режим доступа: [https://www-pub.iaea.org/MTCD/Publications/PDF/P1713\\_R\\_web.pdf](https://www-pub.iaea.org/MTCD/Publications/PDF/P1713_R_web.pdf) [Электронный ресурс]. – (дата обращения: 17.04.2023).

[5] Оценка безопасности установок и деятельности. Нормы безопасности МАГАТЭ. Общие требования безопасности. № GSR Part 4 (Rev. 1). Вена: МАГАТЭ, 2016. – Режим доступа: [https://www-pub.iaea.org/MTCD/Publications/PDF/P1714\\_R\\_web.pdf](https://www-pub.iaea.org/MTCD/Publications/PDF/P1714_R_web.pdf) [Электронный ресурс]. – (дата обращения: 17.04.2023).

[6] Федеральные нормы и правила в области использования атомной энергии «Правила ядерной безопасности реакторных установок атомных станций» (НП-082-07), утвержденные постановлением Федеральной службы по экологическому, технологическому и атомному надзору от 10.12.2007 г. № 4. – Режим доступа: [https://docs.secncs.ru/documents/nps/%D0%9D%D0%9F-082-07/%D0%9D%D0%9F-082-07\\_conv.pdf](https://docs.secncs.ru/documents/nps/%D0%9D%D0%9F-082-07/%D0%9D%D0%9F-082-07_conv.pdf) [Электронный ресурс]. – (дата обращения: 17.04.2023).

## **INFORMATION ENVIRONMENT FOR OPTIMIZATION OF SUPERVISION IN THE FIELD OF NUCLEAR AND RADIATION SAFETY DURING OPERATION OF THE BELARUSIAN NPP**

***S.M. Satsuk***

*Chair of the Department of Electronics of  
BSUIR, PhD, Associate Professor*

***S.V. Drobat***

*Dean of the Faculty of Information Security of  
BSUIR, PhD, Associate Professor*

***U.N. Rusakovich***

*Senior lecturer of the Department of  
Electronics of BSUIR*

***D.O. Serov***

*Student of BSUIR*

*Belarusian State University of Informatics and Radioelectronics, Faculty of Radioengineering and  
Electronics, Republic of Belarus*

*E-mail: satsuk@bsuir.by*

**Abstract.** An information environment has been created for the collection, accumulation, storage and analysis of indicators of safe operation of the Belarusian NPP.

It is shown that the information environment will make it possible to increase the effectiveness of supervisory (control) activities in the field of nuclear and radiation safety in the Republic of Belarus and introduce advanced technologies in the planning and implementation of regulatory activities at all stages of supervision of the Belarusian NPP.

**Keywords:** information environment, nuclear and radiation safety, nuclear power plants, regulatory legal acts.

УДК 621.865.8:629.055

## АЛГОРИТМИКА УПРАВЛЕНИЯ ЧЕТЫРЕХКОЛЕСНЫМИ МОБИЛЬНЫМИ РОБОТАМИ



**С.А. Павлюковец**  
Заведующий кафедры  
«Электропривод и  
автоматизация промышленных  
установок и технологических  
комплексов» БНТУ, кандидат  
технических наук, доцент  
sap@bntu.by



**А.А. Вельченко**  
Доцент кафедры  
«Электропривод и  
автоматизация  
промышленных установок и  
технологических  
комплексов» БНТУ,  
кандидат технических  
наук, доцент  
anna.velchenko@gmail.com



**Ян Шисинь**  
Аспирант кафедры  
«Электропривод и  
автоматизация промышленных  
установок и технологических  
комплексов» БНТУ  
yang\_shi\_xin@163.com



**Д.Ю. Чаплыгин**  
Аспирант кафедры  
«Электропривод и  
автоматизация промышленных  
установок и технологических  
комплексов» БНТУ  
dmitrij.tchapygin@yandex.ru



**А.А. Радкевич**  
студент кафедры  
«Электропривод и  
автоматизация  
промышленных установок и  
технологических  
комплексов» БНТУ  
artyomradkevichbntu@gmail.com



**Н.О. Савко**  
студент кафедры  
«Электропривод и  
автоматизация промышленных  
установок и технологических  
комплексов» БНТУ  
nikitsauko@mail.ru

### **С.А. Павлюковец**

Кандидат технических наук, доцент, заведующий кафедрой «Электропривод и автоматизация промышленных установок и технологических комплексов» Белорусского национального технического университета. Область научных интересов связана с разработкой методов и алгоритмов управления мобильными роботами.

### **А.А. Вельченко**

Кандидат технических наук, доцент, доцент кафедры «Электропривод и автоматизация промышленных установок и технологических комплексов» Белорусского национального технического университета. Область научных интересов связана с разработкой и исследованием нейросетевых регуляторов для электропривода.

### **Ян Шисинь**

Аспирант Белорусского национального технического университета. Область научных интересов связана с исследованием алгоритмов управления колесными мобильными роботами.

### **Д.Ю. Чаплыгин**

Аспирант Белорусского национального технического университета. Область научных интересов связана с исследованием алгоритмов управления робототехническими комплексами для точного земледелия.

**А.А. Радкевич**

Студент 4-го курса Белорусского национального технического университета. Область научных интересов связана с разработкой систем управления электроприводами.

**Н.О. Савко**

Студент 3-го курса Белорусского национального технического университета. Область научных интересов связана с разработкой математических моделей электропривода.

**Аннотация.** В данной работе представлен анализ построения кинематических и динамических моделей четырехколесных мобильных роботов. Для трех случаев передачи вращательного момента (при задней ведущей оси, при передней ведущей оси и при передаче движения к обеим осям мобильного робота), которые определяются характерной структурой уравнений модели, изучены следующие вопросы: голономность и неголономность системы; подвижность и управляемость; конфигурация и механизация (оснащение приводными блоками).

**Ключевые слова:** четырехколесный мобильный робот, неголономность, управление, математическая модель.

### **Введение.**

В современном мире способы передвижения наземных робототехнических систем и комплексов различного назначения характеризуются большим разнообразием. Основой любого колесного мобильного робота (КМР) является шасси той или иной конструкции. Шасси мобильного робота представляет собой совокупность частей, обеспечивающих передачу механической энергии от двигателей к активным элементам движителя – ведущим колесам, звездочкам, шкивам, выходным звеньям механизмов изменения геометрии шасси или механизмов шагания и т.п., – объединенных вместе с электроприводами в функциональную конструктивную подсистему.

Цель настоящей работы заключается в обобщении и представлении общей и единой информации обо всех математических моделях четырехколесных мобильных роботов.

### **Кинематика и динамика мобильного робота.**

Колесный мобильный робот является автоматической интеллектуальной технической системой, выполняющей заданные действия согласно интегрированной в нее базе знаний, перемещающейся в пространстве посредством колес по заданному алгоритму либо самостоятельно определяя траекторию перемещения.

По числу колес выделяют одноколесные, двухколесные, трехколесные и четырехколесные мобильные роботы. По типу колес различают традиционные и всенаправленные (*omni-*) колеса. В первом случае движение колеса по плоскости рассматривается как качение без проскальзывания. При этом скорость точки колеса, в определенный момент времени взаимодействующей с плоскостью движения, равна нулю, а составляющие скорости этой точки колеса равны нулю как в проекции на вектор, лежащий в плоскости колеса, так и в проекции на вектор, перпендикулярный плоскости колеса. В *omni*-колесе лишь одна компонента скорости точки соприкосновения колеса с плоскостью движения равна нулю при движении.

В данной работе описаны математические модели управления применительно к мобильным роботам с колесами традиционного типа для случаев управления различными осями колес: в первом случае ведущая ось – задняя, во втором случае ведущая ось – передняя, в третьем случае обе оси являются ведущими. Примем допущения, что в процессе движения плоскость каждого колеса остается вертикальной по отношению к плоскости движения, и вращение каждого колеса происходит вокруг соответствующей горизонтальной оси, проходящей через центр колеса, а сами колеса и корпус КМР не подвергаются деформации.

Все имеющиеся математические модели управления КМР подразделяются на динамические и кинематические. Динамическая модель описывает движение системы в момент действия на нее внешних сил. Эта модель описывает взаимодействие физических параметров движения и включает параметры энергии, массы системы, инерции и скорости. Описание динамических моделей производится дифференциальными уравнениями второго порядка. Кинематическая модель описывает геометрические взаимодействия, которые присутствуют в

системе. Она описывает взаимосвязь между входными (управляющими) параметрами и реакцией системы, заданной представлением в пространстве состояний. Кинематическая модель описывает быстродействие системы и представляется системой дифференциальных уравнений первого порядка. Для разработки алгоритмов управления КМР обычно достаточно кинематических моделей управления, в то время как для других робототехнических систем, применяемых в космонавтике, авиации или шагающих роботах, также необходимо динамическое моделирование.

Существует несколько типов кинематических моделей:

- внутренняя кинематическая модель описывает взаимосвязь между внутренними переменными системы, например, вращением колеса и движением робота;
- внешняя кинематическая модель описывает положение и ориентацию робота в соответствии с некоторой системой координат;
- прямая кинематика и инверсная кинематика.

Прямая кинематика описывает состояния робота в зависимости от его входных данных (скорости вращения колес, движения суставов, рулевого управления и др.). Инверсная кинематика является способом планирования движения. Это означает, что входные параметры робота могут быть рассчитаны для желаемой последовательности выходных состояний.

Внутренняя кинематическая модель КМР, описывающая его положение на плоскости, определяется трехкомпонентным вектором положения  $q(t)$

$$q(t) = \begin{bmatrix} x_1(t) \\ y_1(t) \\ \varphi(t) \end{bmatrix}, \quad (1)$$

где  $X_1, Y_1$  – координаты точки O центра робота в неподвижной системе координат  $(\vec{X}, \vec{Y})$ ;  
 $\varphi$  – угол между системами координат, определяющими положение подвижного базиса  $(\vec{\alpha}, \vec{\beta})$  относительно неподвижного базиса  $(\vec{X}, \vec{Y})$

Положение четырехколесного мобильного робота на плоскости в этом случае проиллюстрировано на рисунке 1.

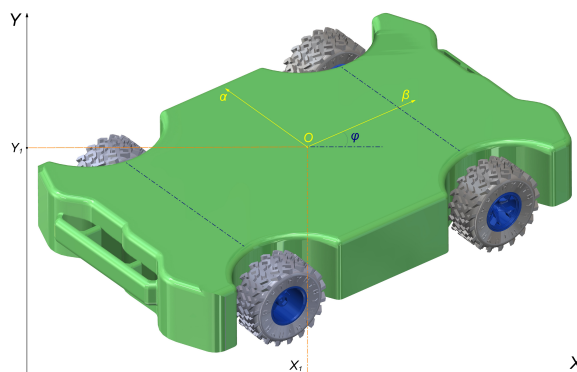


Рисунок 1. Определение положения КМР на плоскости



Поворот подвижной системы координат  $(\vec{\alpha}, \vec{\beta})$  относительно неподвижной системы координат  $(\vec{X}, \vec{Y})$  во внешней кинематической модели определяется транспонированным вектором  $[X_1, Y_1]^T$  и ортогональной матрицей

$$R(\varphi) = \begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

Далее рассмотрим математические модели управления КМР для различных типов приводных осей. Принимаем во внимание, что движение робота осуществляется по горизонтальной плоскости без проскальзывания, и вращение каждого колеса происходит вокруг соответствующей горизонтальной оси, проходящей через центр колеса.

#### Четырехколесный мобильный робот с задней приводной осью.

В четырехколесных мобильных роботах, где все колеса традиционного типа, реализован принцип геометрии рулевого управления Аккермана для поворота передней оси. Он состоит в том, чтобы оси всех колес представлялись в виде радиусов окружностей с общей центральной точкой. Поскольку задние колеса зафиксированы, эта центральная точка должна находиться на линии, продолжающейся от задней оси. Пересечение осей передних колес на этой линии также требует, чтобы одно из колес было повернуто при повороте на больший угол, чем другое колесо.

Кинематическая модель четырехколесного мобильного робота с двумя ведущими задними и двумя передними рулевыми колесами традиционного типа представлена на рисунке 2.

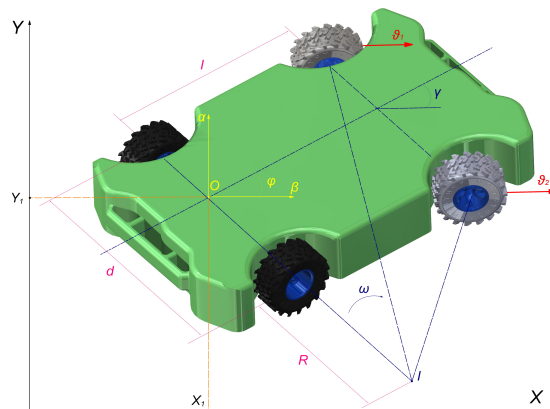


Рисунок 2. Кинематическая модель КМР с задней приводной осью

В качестве переменных состояния колесного робота рассматриваются величины:

$X_1, Y_1$  – координаты базовой точки робота;

$\gamma_1, \gamma_2$  – угол поворота переднего левого и переднего правого колеса соответственно;

$\mathcal{Y}$  – угол поворота платформы робота;

$\varphi$  – угол поворота центральной линии платформы робота;

$V_1, V_2$  – модули вектора скорости правого и левого колеса соответственно;

$L$  – расстояние от базовой точки до передней оси;

$d$  – ширина платформы робота;  
 $I$  – точка мгновенного центра скоростей;  
 $R$  – расстояние от ведущей оси до точки  $I$ ;  
 $\omega$  – угловая скорость вращения ведущей оси.

Направление вращения колес передней оси определяется из выражений

$$\begin{cases} \tan\left(\frac{\pi}{2} - \gamma_1\right) = \frac{R + \frac{d}{2}}{L}; \\ \tan\left(\frac{\pi}{2} - \gamma_2\right) = \frac{R - \frac{d}{2}}{L}. \end{cases} \quad (3)$$

Углы поворота колес передней оси выражаются из уравнений (3)

$$\begin{cases} \gamma_1 = \frac{\pi}{2} - \arctan\left(\frac{R + \frac{d}{2}}{L}\right); \\ \gamma_2 = \frac{\pi}{2} - \arctan\left(\frac{R - \frac{d}{2}}{L}\right). \end{cases} \quad (4)$$

Поскольку правое и левое передние колеса вращаются вокруг мгновенного центра скоростей с одинаковой угловой скоростью  $\omega$ , их окружные скорости равняются

$$\begin{cases} v_1 = \omega \cdot \left(R + \frac{d}{2}\right); \\ v_2 = \omega \cdot \left(R - \frac{d}{2}\right). \end{cases} \quad (5)$$

Кинематическая модель робота описывается следующей системой уравнений

$$\begin{cases} \dot{X} = v \cdot \cos \varphi; \\ \dot{Y} = v \cdot \sin \varphi; \\ \dot{\varphi} = \omega = \frac{v \cdot \operatorname{tg} \gamma}{L}. \end{cases} \quad (6)$$

Перемещение робота описывается системой линейных дифференциальных уравнений

$$\dot{X} = Ax + Bu, \quad (7)$$

где  $A$  и  $B$  – матрицы состояния и управления соответственно;  
 $x$  – вектор состояний;  
 $u$  – вектор управления.

Представляя линейную непрерывную траекторию движения робота в дискретном виде, причем время дискретизации  $\Delta t$  стремится к нулю, а траектория на каждом из участков линейна, запишем систему линейных уравнений (7) в виде

$$X^{K+1} = \tilde{A}x^K + \tilde{B}u, \quad (8)$$

где  $\tilde{A} = A \cdot \Delta t + \varepsilon$ ,  $\tilde{B} = B \cdot \Delta t$ ;

$\varepsilon$  – отклонение от заданной траектории (ошибка позиционирования);

$k$  – шаг дискретизации.

Уравнение (7) возможно представить в матричном виде

$$\begin{bmatrix} v_X^{K+1} \\ v_Y^{K+1} \\ \omega^{K+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} v_X^K \\ v_Y^K \\ \omega^K \end{bmatrix} + B \cdot Q \cdot \Delta t \cdot \begin{bmatrix} \Delta \omega_1 \\ \Delta \omega_2 \\ \Delta \omega_3 \\ \Delta \omega_4 \end{bmatrix}, \quad (9)$$

где матрица управления  $B$  принимает значение

$$B = 2\pi r \cdot \begin{bmatrix} 0,25 & 0,25 & 0,25 & 0,25 \\ -0,25 & 0,25 & 0,25 & -0,25 \\ \frac{1}{2 \cdot (d+l)} & \frac{1}{2 \cdot (d+l)} & -\frac{1}{2 \cdot (d+l)} & \frac{1}{2 \cdot (d+l)} \end{bmatrix}, \quad (10)$$

а матрица поворота  $Q$  имеет вид

$$Q = \begin{bmatrix} \cos \varphi^K & -\sin \varphi^K & 0 \\ \sin \varphi^K & \cos \varphi^K & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (11)$$

### Четырехколесный мобильный робот с передней приводной осью.

Далее рассмотрим математические модели четырехколесного мобильного робота с ведущими передними колесами традиционного типа. Его кинематическая схема отображена на рисунке 3.

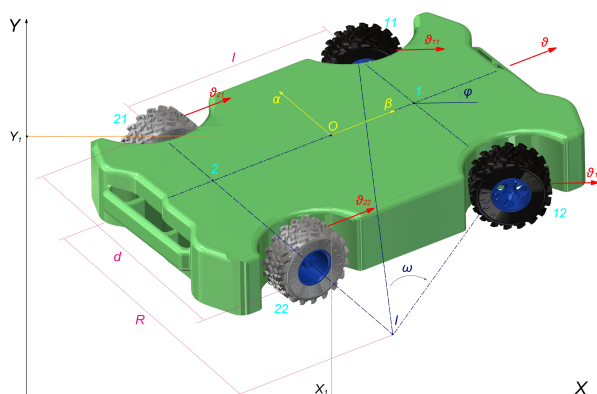


Рисунок 3. Кинематическая модель КМР с передней приводной осью

Данный тип робота использует для движения два передних колеса: 11 – переднее левое колесо, линейная скорость которого обозначена  $g_{11}$ , 12 – переднее правое колесо, линейная скорость которого обозначена  $g_{12}$ , 21 – заднее левое колесо, линейная скорость которого обозначена  $g_{21}$ , 22 – заднее правое колесо, линейная скорость которого обозначена  $g_{22}$ .

Так как мобильный робот оснащен дифференциальным приводом, контроль маневрирования устройства осуществляется путем увеличения или уменьшения скорости на двигателе, приводящем в движение ось. Следовательно, для четырехколесного мобильного робота, линейная скорость переднего привода определяется по формуле

$$g = \frac{g_{11} + g_{12}}{2}. (12)$$

Управление роботом осуществляется с помощью контроля скорости колес, приводящихся в движение двигателями постоянного тока с независимым возбуждением, математическая модель которых имеет следующий вид

$$\begin{cases} \frac{dI_{я}}{dt} = \frac{U - C \cdot \omega_d - I_{я} \cdot R}{L}, \\ \frac{d\omega_d}{dt} = \frac{C \cdot I_{я} - M_C}{J_{дв}}, \end{cases} (13)$$

- где  $I_{я}$  – ток якоря двигателя, А;
- $U$  – напряжение двигателя, В;
- $C$  – конструктивная постоянная двигателя, В·с/рад;
- $\omega_d$  – угловая скорость приводной оси, рад/с;
- $R$  – суммарное сопротивление двигателя, Ом;
- $M_C$  – статический момент на валу двигателя, Н·м;
- $J_{дв}$  – момент инерции двигателя, кг·м<sup>2</sup>.

Тогда математическая модель дифференциальных приводов передней оси может быть записана согласно выражению:

$$\left\{ \begin{array}{l} \frac{dI_{Я11}}{dt} = \frac{U - C \cdot \omega_{11} - I_{Я11} \cdot R}{L}; \\ \frac{d\omega_{11}}{dt} = \frac{C \cdot I_{Я11} - M_C}{J_{ДВ}}; \\ \frac{dI_{Я12}}{dt} = \frac{U - C \cdot \omega_{12} - I_{Я12} \cdot R}{L}; \\ \frac{d\omega_{12}}{dt} = \frac{C \cdot I_{Я12} - M_C}{J_{ДВ}}; \\ \mathcal{G} = \frac{\omega_{11} \cdot R + \omega_{12} \cdot R}{2}, \end{array} \right. (14)$$

где  $\omega_{11}$ ,  $\omega_{12}$  – угловая скорость переднего левого и переднего правого колес соответственно, рад/с;

Кинематическая модель четырехколесного мобильного робота с передней приводной осью имеет следующий вид

$$\left\{ \begin{array}{l} \frac{d\alpha}{dt} = \left( \frac{\mathcal{G}_{11} + \mathcal{G}_{12}}{2} \right) \cdot \sin \varphi + x_0; \\ \frac{d\beta}{dt} = \left( \frac{\mathcal{G}_{11} + \mathcal{G}_{12}}{2} \right) \cdot \cos \varphi + y_0; \\ \frac{d\varphi}{dt} = \left( \frac{\mathcal{G}_{12} - \mathcal{G}_{11}}{L} \right) + \varphi_0, \end{array} \right. (15)$$

где  $\alpha$ ,  $\beta$  – координаты КМР в подвижной системе отсчета;

$\mathcal{G}_{11}$ ,  $\mathcal{G}_{12}$  – линейные скорости левого переднего и правого переднего приводного колеса соответственно;

$L$  – расстояние между осями КМР;

$\varphi$  – угол поворота КМР в подвижной системе координат;

$[x_0, y_0, \varphi_0]$  – начальные координаты в неподвижной системе отсчета, равные нулю.

На основании совмещений математических моделей (14) и (15) получим обобщенную кинематическую модель КМР в виде системы уравнений

$$\left\{ \begin{array}{l} \frac{dI_{я11}}{dt} = \frac{U - C \cdot \omega_{11} - I_{я11} \cdot R}{L}; \\ \frac{d\omega_{11}}{dt} = \frac{C \cdot I_{я11} - M_C}{J_{ДВ}}; \\ \frac{dI_{я12}}{dt} = \frac{U - C \cdot \omega_{12} - I_{я12} \cdot R}{L}; \\ \frac{d\omega_{12}}{dt} = \frac{C \cdot I_{я12} - M_C}{J_{ДВ}}; \\ g = \frac{\omega_{11} \cdot R + \omega_{12} \cdot R}{2}; \\ \frac{d\alpha}{dt} = \left( \frac{g_{11} + g_{12}}{2} \right) \cdot \sin \varphi + x_0; \\ \frac{d\beta}{dt} = \left( \frac{g_{11} + g_{12}}{2} \right) \cdot \cos \varphi + y_0; \\ \frac{d\varphi}{dt} = \left( \frac{g_{12} - g_{11}}{L} \right) + \varphi_0. \end{array} \right. \quad (16)$$

#### Четырехколесный мобильный робот с полным приводом.

Кинематическая модель четырехколесного мобильного полноприводного робота изображена на рисунке 4.

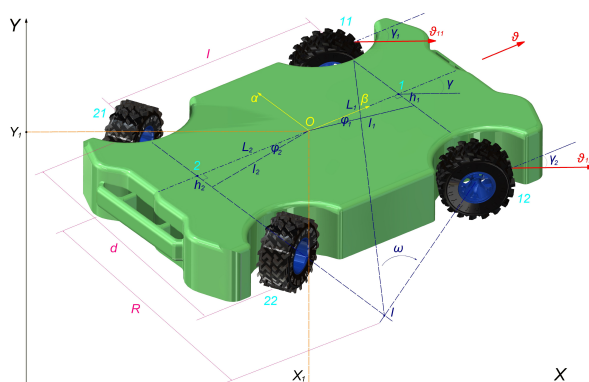


Рисунок 4. Кинематическая модель полноприводного КМР

Запишем координаты центров масс колёс во внешней неподвижной системе координат OXYZ для составления кинематической составляющей модели

$$\left\{ \begin{array}{l} X_{ij} = \alpha + (l_j \cdot \cos \gamma - h_j \cdot \sin \gamma) - l_{ij} \cdot \sin(\gamma + \varphi_i); \\ Y_{ij} = \beta + (l_j \cdot \sin \gamma + h_j \cdot \cos \gamma) + l_{ij} \cdot \cos(\gamma + \varphi_i), \end{array} \right. \quad (17)$$

где  $i$  – номер поворотного блока;

$j$  – номер колеса  $i$ -го поворотного блока;

$l_j, h_j$  – проекции вектора  $L_j$ , соединяющего центр симметрии корпуса робота с точкой крепления  $i$ -го поворотного блока, на оси подвижной системы координат;

$l_{ij}$  – расстояние от точки крепления  $i$ -го поворотного блока до центра  $ij$ -го колеса;

$\varphi_i$  – угол положения  $i$ -го приводного блока в подвижной системе координат.

Продифференцировав по времени уравнения (17), получим выражения для проекций скоростей точек контакта колёс с поверхностью

$$\begin{cases} \mathcal{G}_{Lj}^X = \mathcal{G}_0^X - \omega_0 \cdot (l_l \cdot \sin \gamma + h_l \cdot \cos \gamma) - (\omega_0 + \omega_l) \cdot l_{Lj} \cdot \cos(\gamma + \varphi_l); \\ \mathcal{G}_{Lj}^Y = \mathcal{G}_0^Y + \omega_0 \cdot (l_l \cdot \cos \gamma - h_l \cdot \sin \gamma) - (\omega_0 + \omega_l) \cdot l_{Lj} \cdot \sin(\gamma + \varphi_l), \end{cases} \quad (18)$$

где  $\omega_0$  – угловая скорость корпуса КМР относительно подвижной системы координат;  
 $\omega_l$  – угловая скорость вращения  $l$ -го поворотного блока относительно точки крепления;  
 $\mathcal{G}_0$  – проекции скорости центра корпуса на оси в неподвижной системе координат;  
 $\mathcal{G}_{Lj}$  – проекции скорости  $ij$ -го колеса на оси в неподвижной системе координат.

Приводная сила  $F_{ПП}$  возникает за счёт работы электродвигателя, подключённого к валу колеса. В данном мобильном роботе используются электродвигатели постоянного тока с постоянными магнитами. Математическое описание электропривода  $ij$ -го колеса можно привести к системе дифференциальных уравнений первого порядка (без учёта контуров положения и скорости)

$$\begin{cases} \omega_{ДВЛj} = \frac{\mathcal{G}_{Lj}^{X_{Lj}}}{I_{ЯЛj} \cdot R_{КЛj}}; \\ \frac{d}{dt} \cdot I_{ЯЛj} = \frac{1}{I_{ЯЛj}} \cdot (U_{ФЛj} - K_{ЕЛj} \cdot \omega_{ДВЛj} - R_{ЯЛj} \cdot I_{ЯЛj}); \\ M_{ВЫХ.Д.Лj} = \frac{K \cdot M_{Лj}}{I_{РЛj} \cdot \eta_{РЛj}} \cdot I_{ЯЛj}, \end{cases} \quad (19)$$

где  $\omega_{ДВЛj}$  – угловая скорость вращения электродвигателя  $ij$ -го колеса;

$\mathcal{G}_{Lj}^{X_{Lj}}$  – проекция линейной скорости  $ij$ -го колеса на ось  $O_{Lj}X_{Lj}$  локальной системы координат;

$I_{ЯЛj}$  – ток якоря двигателя;

$R_{ЯЛj}$  – сопротивление обмоток якоря;

$K_{ЕЛj}$  – конструктивная постоянная двигателя;

$M_{ВЫХ.Д.Лj}$  – динамический момент на выходе редуктора;

$I_{РЛj}$  – передаточное число редуктора;

$\eta_{РЛj}$  – коэффициент полезного действия редуктора.

Исходя из рассмотренных выше моделей видно, что алгоритмика управления четырехколесными мобильными роботами представлена кинематическими и динамическими моделями движения робота для трех случаев передачи вращательного момента: при задней ведущей оси, при передней ведущей оси и при передаче движения к обеим осям робота.

### Заключение.

В результате проведенного аналитического исследования получены и структурированы математические модели четырехколесных мобильных роботов, для которых указаны варианты привода и рулевого управления. Для каждого варианта приведены примеры роботов. Проведенная оценка конкретных типов КРМ является качественной и основана на профессиональном опыте авторов и подкреплена доступной литературой. Представленный обзор дает возможность, в соответствии со структурой модели мобильного робота, производить



численные эксперименты и встраивать в систему управления всевозможные модули (датчики), что позволяет получать различные характеристики мобильного робота. Подводя итог, можно сделать вывод, что с учетом всех проанализированных особенностей КМР не существует единого идеального решения во всех аспектах. Выбор кинематической конструкции должен основываться на предполагаемом применении робота.

### Список литературы

- [1] Васильев, А.В. Принципы построения и классификация мобильных роботов наземного применения и планетоходов / А.В. Васильев // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика, телекоммуникации и управление. – 2013. – № 1. – С. 124–131.
- [2] Кампион, Г. Структурные свойства и классификация кинематических и динамических моделей колесных мобильных роботов / Г. Кампион, Ж. Бастен, Б. д'Андреа-Новель // Нелинейная динамика. – 2011. – Т. 7, № 4. – С. 733–769.
- [3] Klančar, Gregor. Wheeled Mobile Robotics. From Fundamentals Towards Autonomous Systems / Gregor Klančar, Andrej Zdešar, Sašo Blažič, Igor Škrjanc. – The Boulevard, Langford Lane, Kidlington, Oxford. – 2017. – 492 p.
- [4] Tzafestas, Spyros G. Mobile Robot Control and Navigation: A Global Overview / Spyros G. Tzafestas // Journal of Intelligent & Robotic Systems. – 2018. – № 91. – P. 35–58.
- [5] Evstigneev, M.I. Algorithms of control over four-wheel robot moving over rough terrain / M.I. Evstigneev, Yu.V. Litvinov, V.V. Mazulina, G.M. Mishchenko // Izvestiya Vysshikh Uchebnykh Zavedeniy. Priborostroenie. – 2015. – Vol. 58, № 9. – P. 738–741 (in Russian).
- [6] Meshkovskiy, E.O. Construction of a mathematical model of a four-wheel mobile robot with two differential drive units / E.O. Meshkovskiy, A.D. Kurmashev // Peter the Great St. Petersburg Polytechnic University. – «Инновации и инвестиции». – 2020. – № 2. – С. 113–118.
- [7] Круглова, Т.Н. Моделирование системы управления полноприводным четырехколесным сельскохозяйственным мобильным роботом / Т.Н. Круглова, А.С. Власов // Вестник БГТУ им. В.Г. Шухова. – 2019. – № 5. – С. 147–154.

## CONTROL ALGORITHM FOR FOUR-WHEEL MOBILE ROBOTS

**S.A. Pauliukavets**

*Head of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU, PhD of Technical Sciences, Associate Professor*

**A.A. Velchenko**

*Associate Professor of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU, PhD of Technical Sciences, Associate Professor*

**Yang Shixin**

*PhD student of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU*

**D.Yu. Tschaplygin**

*PhD student of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU*

**A.A. Radkevich**

*Student of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU*

**N.O. Sauko**

*Student of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU*

*Department of «Electric drive and automation of industrial installations and technological complexes»  
Faculty of Information Technology and Robotics  
Belarusian National Technical University, Republic of Belarus  
E-mail: sap@bntu.by*

**Abstract.** This paper presents an analysis of the construction of kinematic and dynamic models of four-wheeled mobile robots. For three cases of torque transmission (with the rear driving axle, with the front driving axle and with the transmission of motion to both axes of the mobile robot), which are determined by the characteristic structure of the model equations, the following questions were studied: holonomy and nonholonomy of the system; mobility and controllability; configuration and mechanization (equipment with drive units).

**Keywords:** four-wheeled mobile robot, nonholonomy, control, mathematical model.

УДК 004.42

## ИСПОЛЬЗОВАНИЕ BIG DATA В ГОСТИНИЧНОМ БИЗНЕСЕ



**Ю.П. Плахотникова**  
студент БГУИР  
pocketickchaya@gmail.com



**К.В. Спасивцев**  
студент БГУИР  
ksby8819@gmail.com,



**С.Н. Нестеренков**  
Кандидат технических наук, доцент,  
декан факультета  
компьютерных систем и сетей  
s.nesterenkov@bsuir.by

### **Ю.П. Плахотникова**

Студент 4 курса специальности “Информатика и технологии программирования” БГУИР.

### **К.В. Спасивцев**

Студент 4 курса специальности “Информатика и технологии программирования” БГУИР.

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

**Аннотация.** В статье представлен теоретический анализ целесообразности использования больших данных в гостиничном бизнесе. Рассматриваются преимущества, которые предоставляет анализ данных в данной области, а так же возможные риски при его использовании.

**Ключевые слова:** Big Data, гостиничный бизнес

В современном мире продукты и услуги предлагаются в огромных объемах, и каждый из них должен соответствовать определенным потребностям и ожиданиям покупателей. Одним из лучших способов изучить реальные потребности клиентов является анализ данных.

Гостиничный бизнес позволяет реализовать множество возможностей, которые предоставляют данные. Анализ данных может помочь компаниям на всех этапах их деятельности – от разработки бизнес-стратегии до управления компанией и обслуживания клиентов.

Анализ данных помогает компаниям оценить рынок и понять, какие тренды наиболее сильно влияют на данную сферу бизнеса. На основании имеющихся данных можно принимать решения о том, какие услуги и продукты необходимо предложить клиентам в будущем и что необходимо изменить.

Возможных вариантов использования Big Data в гостиничном бизнесе немало. При этом можно выделить следующие основные направления:

1. Улучшение качества обслуживания клиентов. Анализ данных клиентов позволяет понимать их потребности и ожидания, чтобы предложить более качественный сервис

2. Улучшение управления внутренними процессами. Это позволяет отелям оптимизировать и автоматизировать свои бизнес-процессы для более эффективного ведения бизнеса

3. Улучшение эффективности маркетинга. Позволяет получить более точное представление о наиболее эффективных каналах маркетинга, чтобы создавать более эффективные маркетинговые кампании, и тем самым привлекать новых клиентов.

4. Управление стоимостью и ценообразование. Это может помочь отелям лучше понимать свои затраты на услуги и удобства, чтобы оптимизировать свои бюджеты и управлять своими затратами более эффективно.

Ключ к успеху в гостиничном бизнесе – это четкое понимание потребностей и предпочтений клиентов. Анализ данных позволяет определить, какие факторы являются ключевыми для клиентов. Немаловажная информация о гостях может быть получена из данных о бронированиях. Например, при анализе приоритетов гостей при выборе номера, таких как частота их повторных бронирований и среднее время пребывания, можно выявить их индивидуальные предпочтения и найти в них закономерности, применительно к конкретному бизнесу. Из этого можно сделать вывод, какие типы номеров наиболее популярны у гостей и в какое время года отель пользуется наибольшей популярностью. Что позволяет более эффективно использовать свои ресурсы для наиболее полного удовлетворения потребностей гостей. К примеру, если управляете отелем в горах, можно обнаружить, что гости предпочитают наиболее расположенные номера на верхних этажах с панорамным видом, что может стать дополнительным источником дохода. Кроме того, менеджер получает возможность собирать большие объемы данных о предпочтениях гостей в отношении некоторых удобств, таких как, например, спа-залы, бассейн, бани, прочие дополнительные удобства и услуги, которые может предоставлять отель. Все это поможет сконцентрировать свои маркетинговые усилия и напоминать гостям об удобствах проживания в отеле [1].

Гостиничный бизнес должен быть структурирован и оптимизирован, чтобы обеспечить высокий уровень качества обслуживания для клиентов. Управление внутренними бизнес-процессами является центральным для данного типа бизнеса. Сюда можно отнести оптимизацию цепочки логистических поставок, а так же управление внутренними системами отеля, таких как система отопления и кондиционирования. Сбор и анализ данных может оказаться полезными и для оптимизации работы персонала. Эти данные помогут гостиницам установить, какой персонал необходим, чтобы обеспечить качественный сервис.

С использованием Big Data возможен и более эффективный маркетинг. Отели могут использовать данные о своих клиентах, например, об оптимальной длительности пребывания, ценовых предпочтениях, их месте жительства и т.д., чтобы создать более эффективные маркетинговые и рекламные кампании. Сюда же можно отнести ведение и мониторинг социальных сетей. Социальные сети создают возможность для компании информировать клиентов об освобождении номеров до их официальных сообщений, изучать отзывы клиентов и использовать эти данные в будущем для улучшения своих услуг. Крайне полезной является возможность более точно измерить достаточность рекламной кампании в социальных медиа. Также в маркетинге немаловажное значение имеет анализ данных о конкурентах. Бизнес может использовать данные о конкурирующих компаниях, их ценах, маркетинговых и рекламных программах, чтобы понимать, как можно улучшить свои бизнес-процессы и оставаться конкурентоспособными на рынке [2].

Применение Big Data в финансах позволяет бизнесу более эффективно управлять своими финансовыми ресурсами. Данные можно использовать для прогнозирования спроса на гостиничные услуги, определения оптимального уровня цен, а также для предотвращения мошенничества и уменьшения финансовых потерь. Все это позволяет бизнесу экономить деньги, что снижает продолжительность цикла инвестирования и оптимизирует ресурсы для улучшения качества услуг.

Все вышеперечисленное можно отнести к возможностям, которые предоставляет анализ данных гостиничному бизнесу. Однако следует понимать и возможные проблемы и риски, связанные с его использованием:

1. Затраты на обработку данных. Хотя использование Big Data может помочь гостиницам увеличить прибыль и конкурентоспособность, обработка такого объема данных может быть дорогостоящей. Стоимость хранения данных и найм квалифицированных специалистов может стать огромной проблемой для некоторых гостиниц.

2. Безопасность. Сбор информации о клиентах из разных источников может представлять угрозу для их конфиденциальности и безопасности данных. Гостиницам нужно обеспечить надежную защиту данных для предотвращения утечек и нарушений конфиденциальности своих клиентов.

3. Качество данных. Нет гарантии, что данные, собранные из различных источников, являются точными и полными. Некоторые данные могут быть устаревшими, неполными или даже неверными, что может оказаться препятствием для принятия правильных решений.

В конечном итоге, анализ данных в гостиничном бизнесе открывает владельцам множество возможностей. Он позволяет эффективно принимать решения на основе данных и оптимизировать свои бизнес-процессы, что помогает бизнесу адаптироваться к изменчивости рынка и поддерживать свою конкурентоспособность. Однако, решение об использовании больших данных следует принимать в зависимости от возможностей и потребностей конкретной гостиницы. Ведь использование Big Data может как привести к улучшению клиентского сервиса, так и обернуться слишком большими издержками для компании.

#### Список литературы

[1] Как вы можете использовать большие данные для улучшения вашего отеля [Электронный ресурс]. URL: <https://platforms.su/articles/2641> (Дата обращения: 01.04.2023).

[2] Большие данные – будущее туризма [Электронный ресурс]. URL: <https://www.hse.ru/ma/tourism/news/324106392.html> (Дата обращения: 01.04.2023).

[3] Современный отель [Электронный ресурс]. URL: <http://www.hotelexecutive.ru/article.php?numn=12522> – (Дата обращения: 01.04.2023).

## APPLICATION OF BIG DATA IN THE HOTEL BUSINESS

*Y.P. Plakhotnikova*  
Student of BSUIR

*K.V. Spesivtsev*  
Student of BSUIR

*S.N. Nesterenkov*  
PhD, Associate Professor, Dean of  
the Faculty of Computer Systems  
and Networks

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*  
E-mail: [packetickchaya@gmail.com](mailto:packetickchaya@gmail.com), [ksby8819@gmail.com](mailto:ksby8819@gmail.com), [s.nesterenkov@bsuir.by](mailto:s.nesterenkov@bsuir.by)

**Abstract.** The article presents a theoretical analysis of the use of Big Data in the hotel business. The advantages that are provided with the data analysis are considered. There are also risks associated with its use.

**Keywords:** Big Data, hotel business

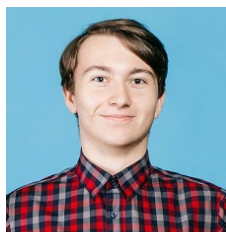
УДК 004.37

## ПРЕИМУЩЕСТВА ИСПОЛЬЗОВАНИЯ BIG DATA В ОБРАЗОВАНИИ И УЧЕБНОМ ПРОЦЕССЕ



**Д.А. Сухарко**

Студент 4 курса, кафедры ПОИТ,  
инженер-программист  
95100239@study.bsuir.by,



**В.Д. Владымицев**

Ассистент кафедры  
информатики, инженер  
программист ОИАСУ ЦИИР  
БГУИР  
v.vladymtsev@bsuir.by,



**А.Н. Марков**

Старший преподаватель,  
заместитель начальника Центра  
информатизации и  
инновационных разработок  
БГУИР  
a.n.markov@bsuir.by

**Д.А. Сухарко**

Студент 4 курса “Программное обеспечение информационных технологий” БГУИР.

**В.Д. Владымицев**

Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР

**А.Н. Марков**

Старший преподаватель кафедры ПОИТ, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Использование Big Data в образовании имеет большой потенциал для улучшения качества обучения и управления учебным процессом. Благодаря сбору и анализу больших объемов данных можно получить ценные знания о процессах обучения и потребностях учащихся, что позволяет более эффективно организовать учебный процесс и персонализировать обучение. Однако, при использовании Big Data в образовании необходимо соблюдать высокие стандарты защиты данных и учитывать этические принципы.

**Ключевые слова:** big data, образование, учебный процесс, управление учебным процессом, предсказание успеваемости, технические аспекты, педагогические аспекты, защита данных, методики обучения, инструменты анализа данных, рекомендации, направления исследований.

### Введение.

В последние годы, использование больших данных (Big Data) становится все более популярным во многих областях, включая образование. Большие данные (Big Data) могут стать мощным инструментом для преобразования формы обучения, переосмысления подходов, сокращения проблем и адаптирования опыта для повышения эффективности самой образовательной системы [1]. В данном тезисе мы рассмотрим преимущества использования Big Data в сфере образования и учебном процессе.

Цель данной работы - проанализировать преимущества использования Big Data в сфере образования и учебном процессе и оценить потенциал данной технологии для улучшения качества обучения. В работе были рассмотрены основные преимущества использования Big Data в образовании, такие как персонализированное обучение, улучшение управления учебным процессом и предсказание успеваемости учащихся. Также были проанализированы основные технические, педагогические и этические аспекты использования Big Data в образовании. В результате работы были выдвинуты рекомендации по использованию Big Data в образовании и указаны основные направления дальнейших исследований в данной области.

### Анализ данных учащихся.

Одним из преимуществ использования Big Data в образовании является возможность

анализировать данные учащихся. Данные могут быть получены из различных источников, таких как учебные материалы, тесты, экзамены и другие формы оценки. Анализ этих данных помогает педагогам и учителям понять, как учащиеся усваивают материал, и определить, какие учебные методы и стратегии работают лучше всего для каждого ученика. Это может привести к более эффективному учебному процессу и повышению успеваемости учащихся [2].

#### **Персонализированное обучение.**

Еще одним преимуществом использования Big Data в образовании является возможность создания персонализированного обучения для каждого ученика. Анализ данных учащихся позволяет педагогам и учителям создавать индивидуальные учебные планы, которые учитывают индивидуальные потребности и способности каждого ученика. Персонализированное обучение приводит к большей вовлечённости учащегося, а учителю помогает понять уровень каждого ученика в отдельности, что помогает обеспечить соответствующее руководство и дополнительные ресурсы для повышения академического уровня учеников [3]. Это повышает мотивацию учащихся и помогает им достичь лучших результатов.

#### **Улучшение управления образовательными учреждениями.**

Big Data также может помочь улучшить управление образовательными учреждениями. Анализ данных может помочь учителям и администрации школы улучшить планирование и распределение ресурсов, что позволит повысить эффективность учебного процесса и сократить затраты на образование.

#### **Прогнозирование успеваемости учащихся.**

Использование Big Data также может помочь прогнозировать успеваемость учащихся на основе анализа данных. Анализ исторических данных и машинное обучение могут использоваться для прогнозирования, какие учащиеся могут испытывать трудности в определенных предметах, и на основе этого педагоги могут предпринимать дополнительные меры, чтобы помочь им улучшить свои знания и навыки [4].

#### **Инновационные методы обучения.**

Big Data также может стимулировать разработку новых, инновационных методов обучения. Использование Big Data позволяет идентифицировать тренды в учебном процессе и определять, какие методы и подходы наиболее эффективны [5]. Это может привести к созданию новых методов обучения, которые лучше подходят для учащихся и позволяют им учиться более эффективно.

В сфере образования выделяются пять основных типов данных: персональные данные; данные о взаимодействии студентов с электронными системами обучения (электронными учебниками, онлайн-курсами); данные об эффективности учебных материалов; административные (общесистемные) данные; прогнозные данные.

На основе анализа множества подходов и моделей в своих выступлениях И.Д. Фрумин выделяет три крупных направления Big Data [6]:

- 1) связанные с мышлением (прежде всего критическим и креативным мышлением);
- 2) связанные со взаимодействием с другими (коммуникация и коллаборация);
- 3) связанные со взаимодействием с самим собой (саморегулирование, рефлексивность и самоорганизация).

Но результаты аналитики в данных направлениях наиболее ценны, когда выявляют аномальные и пограничные состояния образовательной системы. Меры регулирования как реакция на отрицательные состояния наиболее полезны для работы по развитию образовательной системы.

Для того чтобы использование Big Data в образовании было максимально эффективным, необходимо учитывать не только технические аспекты, но и педагогические. Важно разработать новые методики обучения, которые будут основаны на данных, полученных с помощью Big Data. Такие методики могут включать в себя персонализированное обучение, в котором учебный процесс будет адаптироваться под конкретного учащегося, а также использование

инновационных технологий, таких как виртуальная реальность, искусственный интеллект и др.

Важно также понимать, что использование Big Data в образовании не является панацеей и не может заменить человеческий фактор. Педагоги всегда будут играть важную роль в образовании, и использование Big Data должно служить в качестве дополнения к их работе. Кроме того, необходимо учитывать мнение учащихся и принимать во внимание их потребности и интересы, чтобы обеспечить максимальную эффективность образовательного процесса.

#### **Заключение.**

Использование Big Data в образовании имеет огромный потенциал для улучшения качества обучения и управления учебным процессом. Однако, для того чтобы использование Big Data было максимально эффективным, необходимо учитывать технические, педагогические и этические аспекты. Важно также разработать новые методики обучения, которые будут основаны на данных, полученных с помощью Big Data. Использование Big Data в образовании должно служить в качестве дополнения к работе педагогов и учитывать потребности и интересы учащихся.

#### **Список литературы**

- [1] Стивен Дагэн. Искусственный интеллект в образовании: Изменение темпов обучения. Аналитическая записка ИИТО ЮНЕСКО // Институт ЮНЕСКО по информационным технологиям в образовании, 2020.
- [2] Ваганова О.И., Дворникова Е.И., Кутепов М.М., Лунева Ю.Б., Трутанова А.В. Возможности облачных технологий в электронном обучении // Международный журнал прикладных и фундаментальных исследований. – 2017. – № 6-2. – С. 183-187.
- [3] Нестеренков, С.Н. Функциональная модель процедур планирования и управления образовательным процессом как основа построения информационной среды учреждения высшего образования / С.Н. Нестеренков, Н.В. Лапицкая // Вести Института современных знаний. - 2018. - N 1. - С. 97-105.
- [4] Фельдштейн Д. И. Приоритетные направления психолого-педагогических исследований в условиях значимых изменений ребенка и ситуации его развития. – Воронеж; М.: МПСИ, Модэк, 2010. – 16 с.
- [5] Корпоративное обучение для цифрового мира / под ред. В. С. Катькало, Д. Л. Волкова. – М.: АНО ДПО «Корпоративный университет Сбербанка», 2017. – 200 с.:
- [6] Фрумин И. Д. Тренды в развитии содержания образования: ключевые компетенции и новая грамотность // Материалы IV Международного форума по педагогическому образованию. – URL: <http://ifte.kpfu.ru/ru/lectures/trendy-v-razviti-i-sod>

## **THE BENEFITS OF USING BIG DATA IN EDUCATION AND THE LEARNING PROCESS**

***D.A. Sukharko***

*Student of BSUIR, software engineer*

***V. D. Vladymtsev***

*Assistant of the Department of Computer Science, Software Engineer of DIACS CIIR BSUIR*

***A. N. Markov***

*Senior lecturer of the department, Deputy head of the Center for Informatization and Innovative Developments*

*Center for Informatization and Development of the Belarusian University of State Informatics and Radioelectronics, Republic of Belarus*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*

*E-mail: 95100239@study.bsuir.by, v.vladymtsev@bsuir.by, a.n.markov@bsuir.by*

**Abstract.** The use of Big Data in education has great potential for improving the quality of education and managing the educational process. By collecting and analyzing large amounts of data, you can gain valuable knowledge about learning processes and student needs, which allows you to organize the learning process more effectively and personalize learning. However, when using Big Data in education, it is necessary to comply with high data protection standards and take into account ethical principles.

**Keywords:** big data, education, educational process, educational process management, academic performance prediction, technical aspects, pedagogical aspects, data protection, teaching methods, data analysis tools, recommendations, research directions.



УДК 004.043+ 004.67

## ВЕБ-АНАЛИТИКА КАК ИНСТРУМЕНТ ДЛЯ ИЗМЕРЕНИЯ И АНАЛИЗА ЭФФЕКТИВНОСТИ МАРКЕТИНГОВЫХ КАМПАНИЙ



**Д.В. Тавлуй**

Магистрант кафедры проектирования информационно-компьютерных систем БГУИР  
dmitriytavlyuy@gmail.com



**В.Ф. Алексеев**

Доцент кафедры проектирования информационно-компьютерных систем БГУИР, кандидат технических наук, доцент  
alexvikt.minsk@gmail.com



**Г.А. Пискун**

Доцент кафедры проектирования информационно-компьютерных систем БГУИР, кандидат технических наук, доцент  
alexvikt.minsk@gmail.com

### **Д.В. Тавлуй**

Окончил БГУИР (2021 г.), в настоящее время является магистрантом этого университета. Проводит научные исследования по алгоритмам и методам интернет-маркетинга в конкурентной среде.

### **В.Ф. Алексеев**

Окончил Минский радиотехнический институт. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

### **Г.А. Пискун**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с моделированием и оптимальным проектированием информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

**Аннотация.** В статье рассматриваются подходы в реализации рекламной компании в интернете, описываются эффективные решения на каждом из этапов, выделяются наиболее оптимальные методы достижения целей рекламной компании. Показано, что для измерения и анализа эффективности маркетинговых кампаний целесообразно использовать Веб-аналитику.

**Ключевые слова:** рекламная компания, интернет-маркетинг, целевая аудитория, анализ конкурентов, Веб-аналитика.

### **Введение.**

Интернет-маркетинг является одним из ключевых инструментов продвижения продуктов и услуг в сети. Комплексному изучению проблем интернет-маркетинга посвятили свои исследования авторы: Н. Патель, С. Годин, Г. Вэйчжун, Р. Холидей, Д. М. Скотт, Д. Ариели, Э. Рис, Б. Солис, И. Сегалович, А. Шевченко, С. Мироненко, А. Юрьевич, О. Козачук, А. Медина, А. Жук и другие. Следует подчеркнуть, что труды указанных авторов посвящены фундаментальным проблемам. Но множество прикладных вопросов интернет-маркетинга требуют теоретических разработок или дальнейшего совершенствования, особенно в условиях столь стремительного развития интернет-технологий [1–9].

Реклама в интернет-маркетинге представляет собой совокупность мероприятий, направленных на продвижение товаров и услуг в Интернете. Реклама может включать в себя различные форматы, такие как баннеры, контекстную рекламу, видеорекламу, рекламу в социальных сетях и другие.

Одна из главных особенностей рекламы в интернет-маркетинге – это возможность точно настроить таргетирование и доставить рекламу нужной аудитории, что позволяет максимально эффективно использовать бюджет рекламной компании и получить максимальную отдачу от



рекламы [1–7].

### **Реклама в интернете как эффективная форма маркетинга.**

Реклама в интернете — это любая рекламная деятельность, которая использует Интернет для достижения своих целей. Реклама в Интернете стала очень популярной и эффективной формой маркетинга благодаря многим преимуществам, которые она предлагает, таким как:

1. *Широкий охват аудитории.* Интернет — это глобальная платформа, которая позволяет довольно легко достигать аудиторию в разных уголках мира. Рекламные кампании могут быть направлены на конкретную аудиторию, основываясь на интересах и поведении пользователей.

2. *Гибкость.* Реклама в Интернете может быть создана и оптимизирована в режиме реального времени. Это позволяет быстро реагировать на изменения в поведении пользователей и менять рекламные кампании в соответствии с этими изменениями.

3. *Более низкие затраты.* Реклама в Интернете может быть гораздо дешевле, чем традиционные формы рекламы, такие как телевизионная или радиореклама. Это позволяет небольшим компаниям с ограниченным бюджетом участвовать в конкуренции с крупными компаниями.

4. *Высокая измеримость.* Результаты рекламной кампании в Интернете могут быть легко измерены и анализированы. Это позволяет компаниям оценить эффективность своих кампаний и вносить корректировки, если это необходимо.

5. *Разнообразие форматов.* Реклама в Интернете может быть представлена в различных форматах, таких как баннерная реклама, видеореклама, контекстная реклама, реклама в социальных сетях и т.д. Это позволяет компаниям выбрать наиболее подходящий формат для своей рекламной кампании [1].

Однако, реклама в Интернете также может иметь свои недостатки, например, низкую конверсию, риск получения негативной реакции со стороны пользователей, проблемы с подделкой трафика и т.д.

Цели рекламы в интернете могут варьироваться в зависимости от конкретной компании и ее бизнес-стратегии. Основной целью является привлечение новых клиентов, удержание существующих клиентов, увеличение продаж, расширение аудитории, улучшение узнаваемости бренда и продукта, улучшение репутации компании, сбор данных о клиентах, улучшение конверсии [2].

### **Форматы рекламы в интернете.**

Существуют различные форматы рекламы в интернете, которые могут быть использованы для достижения различных целей рекламной кампании. Некоторые из наиболее распространенных форматов рекламы в интернете включают:

1. *Баннерная реклама* – это вид интернет-рекламы, который использует графические изображения (баннеры), размещаемые на веб-сайтах, чтобы привлечь внимание пользователей и направить их на сайт рекламодателя. Баннеры могут иметь различные размеры, формы и содержание, и могут размещаться на разных площадках, таких как сайты, социальные сети, поисковые системы и другие онлайн-сервисы. Баннерная реклама имеет несколько преимуществ по сравнению с другими формами интернет-рекламы. Она позволяет привлечь внимание пользователей, которые просто просматривают сайты, не ищут конкретную информацию. Баннеры могут также использоваться для целевой рекламы, когда рекламодатель стремится достичь определенной аудитории, настроив показ объявлений на определенных сайтах, которые часто посещают пользователи из этой аудитории.

2. *Текстовая реклама* – это вид интернет-рекламы, который использует текстовые объявления, размещаемые на различных площадках в интернете, таких как поисковые системы, социальные сети, сайты и т.д. Текстовые объявления обычно состоят из заголовка, описания и ссылки на сайт рекламодателя. Она позволяет точно нацелиться на целевую аудиторию, используя ключевые слова и фразы, которые пользователи вводят в поисковые системы или

интересуются ими. Текстовая реклама также обычно стоит дешевле, чем другие виды интернет-рекламы, и легче создается и поддерживается.

3. *Реклама в социальных сетях* является одним из наиболее эффективных способов продвижения продуктов и услуг в интернете. Социальные сети, такие как *Facebook, Instagram, Twitter, LinkedIn, TikTok* и другие, предоставляют рекламодателям возможность достигать своей целевой аудитории с помощью точного таргетирования на основе множества параметров, таких как географическое местоположение, возраст, пол, интересы, поведенческие характеристики и многое другое.

4. *Поисковая реклама* – это форма онлайн-рекламы, при которой рекламодатель показывает свою рекламу в результатах поиска на поисковых системах (например, *Google* или *Яндекс*) по определенным ключевым словам или фразам. Данная форма рекламы позволяет рекламодателям привлекать внимание целевой аудитории, которая ищет информацию по конкретным запросам в поисковых системах. Рекламодатель оплачивает каждый клик по его рекламе (*Pay-per-click, PPC*), что означает, что рекламодатель платит только в том случае, если пользователь кликнул на его объявление и перешел на сайт рекламодателя. Поисковая реклама имеет множество преимуществ, включая быстрый старт, гибкость, измеримость и оптимизируемость результатов. Она также позволяет рекламодателям быстро адаптироваться к изменениям в потребительском спросе и менять свою рекламу в соответствии с этими изменениями.

5. *Реклама в приложениях* — это форма онлайн-рекламы, которая показывается в мобильных приложениях на устройствах, таких как смартфоны и планшеты. Рекламодатели могут использовать различные форматы рекламы, такие как баннеры, видеоролики, интерактивные объявления, настраиваемые баннеры и другие форматы.

Реклама в приложениях обычно показывается в определенных местах в приложении, таких как внизу экрана, внутри меню, внутри контента или при входе/выходе из приложения. Она может быть направлена на определенную аудиторию в соответствии с их интересами, поведением и другими параметрами. Данная форма рекламы предоставляет рекламодателям возможность достигнуть широкой аудитории, особенно молодых и мобильных пользователей, которые тратят большую часть своего времени в приложениях на своих мобильных устройствах. Кроме того, реклама в приложениях может быть более эффективной, чем реклама в браузере, поскольку она может быть лучше интегрирована в пользовательский интерфейс приложения и не прерывает поток использования приложения. Однако, реклама в приложениях может вызвать негативную реакцию у пользователей, если она слишком интенсивна или некачественна. Поэтому, рекламодатели должны учитывать предпочтения пользователей и стремиться создавать рекламу, которая является интересной и релевантной для своей аудитории [3].

6. *Реклама в электронной почте*, или *E-mail-маркетинг*, является одним из наиболее распространенных способов продвижения товаров и услуг в интернете. Однако, необходимо учитывать, что многие пользователи рассматривают такую рекламу как навязчивую и могут относиться к ней негативно.

7. *Ретаргетинг* – это метод онлайн-рекламы, при котором рекламные объявления показываются пользователям, которые ранее посещали ваш веб-сайт или взаимодействовали с вашим брендом.

Ретаргетинг позволяет продолжить коммуникацию с теми, кто уже проявил интерес к вашему продукту или услуге, и вернуть их на сайт для завершения покупки или выполнения других действий. Процесс ретаргетинга начинается с установки кода отслеживания на ваш сайт. Когда пользователь посещает ваш сайт, код отслеживания сохраняет данные о пользователе, которые могут быть использованы для определения того, какие объявления показывать ему в будущем. Когда этот пользователь затем посещает другие сайты, объявления вашего бренда могут быть показаны ему на этих сайтах.

Ретаргетинг позволяет более эффективно использовать рекламный бюджет, так как объявления показываются только тем, кто уже проявил интерес к вашему бренду. Однако, важно не злоупотреблять этим методом, чтобы не вызывать у пользователей раздражение и не нарушать их конфиденциальность. Кроме того, необходимо обязательно учитывать требования законодательства по защите персональных данных [4].

### Выбор целевой аудитории.

Сегодня интернет-маркетинг является неотъемлемой частью многих бизнес-стратегий и продолжает активно развиваться, включая в себя новые технологии, такие как искусственный интеллект, *Big Data* и машинное обучение.

Интернет-маркетинг обладает рядом преимуществ перед традиционными методами маркетинга. Он характеризуется:

- большей точностью в измерении эффективности кампаний и ROI;
- охватывает широкую аудиторию и дает возможность детального таргетинга;
- имеет низкие затраты на рекламу по сравнению с традиционными методами;
- обладает возможностью быстрого реагирования на изменения в поведении и запросах клиентов.

Одной из ключевых характеристик интернет-маркетинга является целевая аудитория. Выбор целевой аудитории для рекламы значимый этап, поскольку это поможет определить, кому следует направить свою рекламу, чтобы достигнуть наилучших результатов. Зная целевую аудиторию, можно создавать рекламу, которая наиболее точно соответствует потребностям и интересам клиентов. Это позволит увеличить вероятность того, что они будут заинтересованы в продукте или услуге, и соответственно повысит эффективность рекламы. Тем самым повышается эффективность рекламы. Можно выбрать те каналы маркетинга, которые наиболее подходят для вашей аудитории. Например, если ваша целевая аудитория молодежь, то использование социальных медиа будет наиболее эффективным, чем телевизионная реклама, которая обойдется вам дороже. Это позволит сократить расходы на рекламу и использовать более эффективные каналы маркетинга.

Определение целевой аудитории включает в себя анализ потенциальных клиентов на основе различных характеристик, таких как возраст, пол, доход, образование, интересы, поведение и многое другое. По результатам анализа формируется профиль идеального клиента, который называется «байер персона» (*buyer persona*), структура показана на рисунке 1.

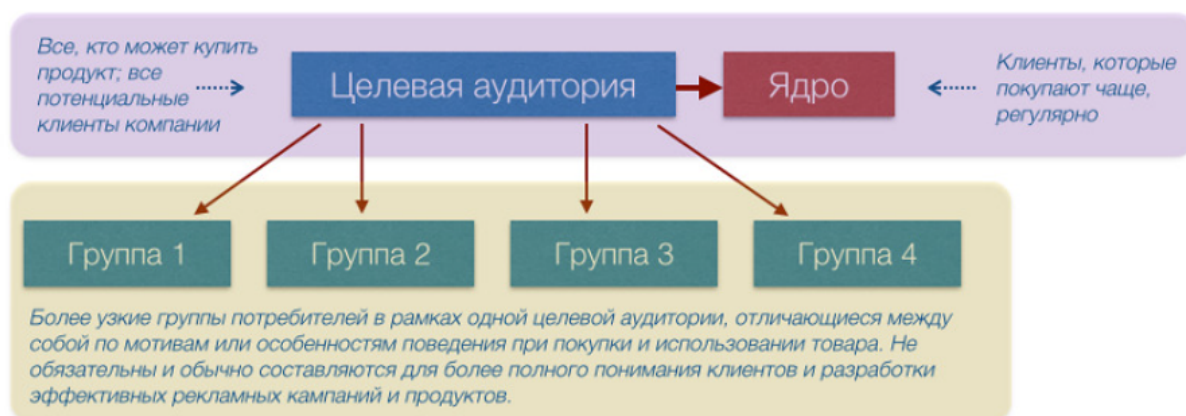


Рисунок 1. Структура целевой аудитории

Целевая аудитория может быть определена для различных маркетинговых кампаний, таких как реклама, E-mail маркетинг, социальные сети, SEO и т.д. Понимание потребностей и

характеристик целевой аудитории помогает компаниям создавать более эффективные и персонализированные маркетинговые кампании, которые приводят к большему количеству продаж и удовлетворенности клиентов [5].

### **Анализ конкурентов.**

Важное значение для повышения эффективности маркетинговых кампаний является анализ рекламной кампании конкурентов. Этот процесс позволяет получить ценную информацию о фирмах-конкурентах, их стратегиях и тактиках в интернет-рекламе. Ниже приведены рекомендуемые основные шаги анализа конкурентов:

1 *Изучение сайта конкурента.* Оцените дизайн сайта, удобство использования, наличие важных функций и технологий, таких как мобильная версия, формы обратной связи, возможность регистрации, онлайн-чата и т.д.

2 *Изучение ключевых слов.* Используйте специальные инструменты для анализа ключевых слов, чтобы выявить, на какие запросы конкуренты нацеливаются в своих рекламных кампаниях.

3 *Изучение рекламных материалов.* Оцените рекламные объявления конкурентов, чтобы определить, какие форматы и подходы они используют, и какие преимущества и уникальные предложения они предлагают.

4 *Изучение социальных сетей.* Изучите социальные сети конкурентов, чтобы понять, как они используют этот канал для привлечения и удержания клиентов.

5 *Изучение отзывов.* Изучите отзывы клиентов о конкурентах на сайтах отзывов и в социальных сетях. Это поможет понять, какие проблемы у конкурентов, и какие преимущества вы можете использовать для своей рекламы.

6 *Анализ бэклинков.* Изучите, какие сайты ссылаются на конкурентов, чтобы понять, как они укрепляют свой авторитет в интернете, и как можно повысить свой собственный авторитет.

После тщательного анализа конкурентов можно разработать свою рекламную стратегию, учитывая преимущества и недостатки конкурентов, и создать рекламные материалы, которые будут эффективными и привлекательными для нашей целевой аудитории [6].

### **Мониторинг и оптимизация результатов рекламы.**

Аналитика является неотъемлемой частью интернет-маркетинга и используется для измерения и анализа эффективности маркетинговых кампаний, а также для принятия решений о дальнейшей стратегии продвижения.

С помощью аналитики можно измерять количество посещений сайта, длительность сессий, конверсию, *ROI (Return on investment – коэффициент рентабельности инвестиций)*, а также другие метрики, которые помогают определить эффективность маркетинговых кампаний и выявить проблемные места в работе сайта [7–9].

Расчет *ROI* необходим, когда нужно узнать, в каком объеме возвращаются средства, вложенные в команду, проект или рекламную кампанию. Этот показатель часто применяют для подсчета окупаемости разных видов рекламы.

Для анализа эффективности маркетинговых кампаний используются различные инструменты, такие как *Google Analytics*, Яндекс.Метрика, *SEMrush*, *Ahrefs* и другие. На рисунке 2 показан дашборд с инструментами аналитики. Эти инструменты помогают собрать данные о посетителях сайта, оценить качество трафика, измерить конверсию и определить места, где можно улучшить работу сайта.

Расчет *ROI* необходим, когда нужно узнать, в каком объеме возвращаются средства, вложенные в команду, проект или рекламную кампанию. Этот показатель часто применяют для подсчета окупаемости разных видов рекламы.

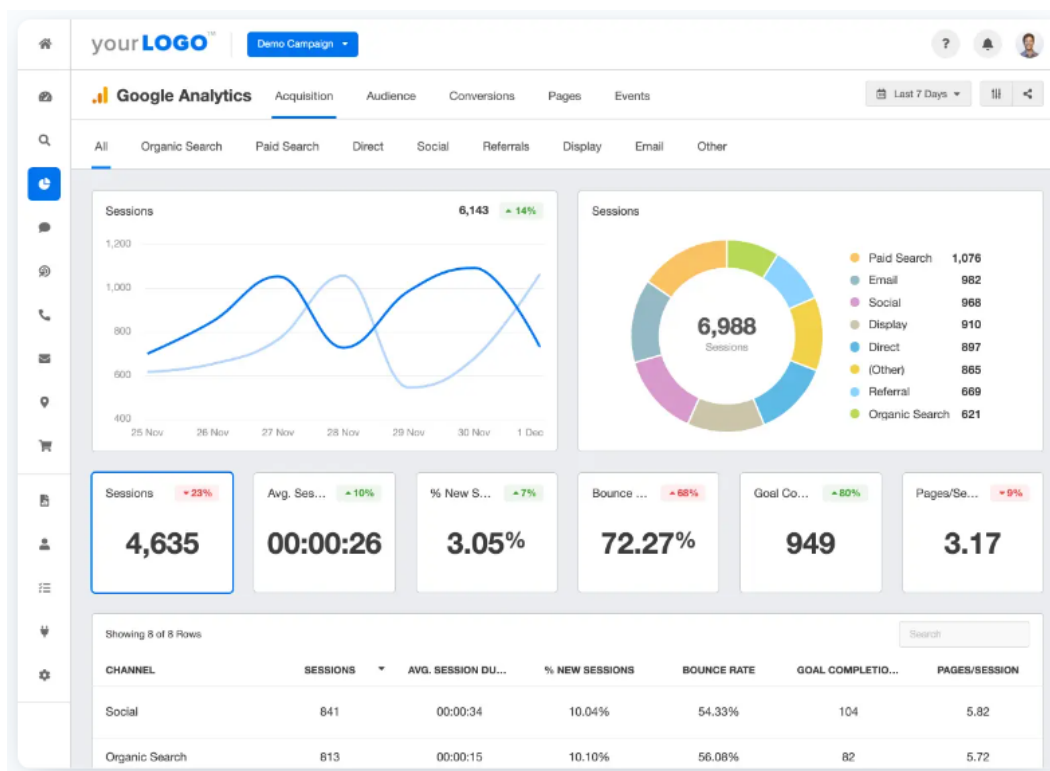


Рисунок 2. Дашборд Google Analytics

Отслеживание ключевых показателей эффективности (KPI). На данном этапе определяется, какие показатели вам важны для оценки успеха рекламной кампании, например, количество кликов, конверсия, средний чек, ROI и другие. Эти KPI помогут понять, насколько хорошо работает рекламная кампания. Также важным шагом является использование аналитических инструментов для отслеживания и анализа данных о посещаемости сайта, поведении пользователей, и других метриках, которые могут помочь определить, какие элементы рекламной кампании работают хорошо, а какие нуждаются в оптимизации. Немаловажным является тестирование и оптимизация. Необходимо тестировать различные варианты рекламных объявлений и лендингов, чтобы определить, какие из них работают лучше. Нужно изменять параметры кампании, такие как таргетинг, ключевые слова, ставки за клик и другие, чтобы улучшить результаты кампании [7–9].

Одним из важных преимуществ аналитики в интернет-маркетинге является возможность быстро реагировать на изменения в поведении и запросах клиентов. На основе данных, собранных в результате анализа, можно изменять стратегию продвижения, улучшать пользовательский опыт на сайте, оптимизировать рекламные кампании. Аналитика в интернет-маркетинге является важным инструментом, который помогает брендам оптимизировать свою стратегию продвижения, повышать эффективность маркетинговых кампаний и увеличивать количество потенциальных клиентов [7-9].

Одним из основных преимуществ интернет-маркетинга является возможность использования более низких затрат на продвижение товаров и услуг в сравнении с традиционными методами маркетинга, такими как телевидение, радио или пресса.

### Заключение.

Рекламная компания в интернете — это один из самых эффективных способов продвижения товаров и услуг на рынке. Важно понимать, что интернет предоставляет огромные возможности для достижения целевой аудитории, благодаря своей глобальной и многоязычной природе. Рекламная компания включает в себя множество шагов и этапов, каждый из которых

может критически повлиять на дальнейшее продвижение. И только эффективное применение методов и подходов на каждом из этапов приведет к достижению целей рекламной компании.

#### **Список литературы**

- [1] Голик, В. Учебно-методический комплекс "Интернет-маркетинг и электронная коммерция" / В. Голик. – Минск : БГУ, 2017. – 101-104 с.
- [2] Тазова, В.Д. Контент-маркетинг, как инструмент продвижения компании в интернете / В.Д. Тазова, И.А. Стоянов. – Красноярск : Красноярский государственный аграрный университет, 2017. – 55-60 с.3. Жлукта, П.А. Эффективность внедрения облачной cтп-системы / П.А. Жлукта. – Минск : Белорусский государственный университет экономического факультета, 2019. – 78-90 с.
- [3] Воронюк, А. Актуальный интернет-маркетинг. Как привлекать клиентов и увеличивать продажи через Интернет / А. Воронюк, А. Полищук. – Киев : Агентство «РІО», 2018. – 24-30 с.
- [4] Алексеев, В.Ф. Разработка онлайн платформы оценки и финансирования инновационных проектов / В.Ф. Алексеев, Д.В. Лихачевский, Г.А. Пискун // BIG DATA Advanced Analytics: collection of materials of the fourth international scientific and practical conference, Minsk, Belarus, May 3 - 4, 2018 / editorial board: M. Batura [etc.]. - Minsk, BSUIR, 2018. - P. 398 – 404.
- [5] Алексеев, В.Ф. Автоматизация стратегического прогнозирования и планирования технико-экономических показателей в маркетинговых исследованиях / В.Ф. Алексеев, В.И. Журавлев, В.И. Лакин // Экономическое развитие общества: инновации, информатизация, системный подход. Международная научно-практическая конференция. Тезисы докладов. - Минск: Изд-во «ПАРАДОКС», 2008. – С. 170 – 173.
- [6] Алексеев, В.Ф. Анализ системы маркетинга на предприятии и её совершенствование с использованием Internet-технологий / В. Ф. Алексеев [и др.] // Современные информационные компьютерные технологии: Сб. науч. ст. в 2ч. Ч.1 - Гродно: ГрГУ, 2008. – С. 118 – 122.
- [7] Алексеев, В.Ф. Задачи и методы обработки социально-экономической информации / В.Ф. Алексеев // Современные средства связи: материалы XVI междунар. науч.-техн. конф., 27–29 сентября 2011 г. — Минск: УО ВГКС, 2011.— С. 102.
- [8] Андреева, О.Д. Технология бизнеса: маркетинг / О.Д. Андреева. – Красноярск : ИНФРА-МНОРМА, 2019. = 456 с.
- [9] Артюхова, Т.З. Особенности маркетинга в интернете на примере российских социальных сетей / Т.З. Артюхова, Т.А. Пырьева. – Красноярск : Инновационная наука, 2017. – 155-157 с.

## **WEB ANALYTICS AS A TOOL FOR MEASURING AND ANALYZING THE EFFICIENCY OF MARKETING CAMPAIGN**

***D.V. Tavlyuy***

*Master student of the department  
design of information and  
computer systems BSUIR*

***V.F. Alexseev***

*Associate Professor of the  
Department of Design of  
Information and Computer  
Systems of BSUIR, PhD of  
Technical Sciences, Associate  
Professor*

***G.A. Piskun***

*Associate Professor of the  
Department of Design of  
Information and Computer  
Systems of BSUIR, PhD of  
Technical Sciences, Associate  
Professor*

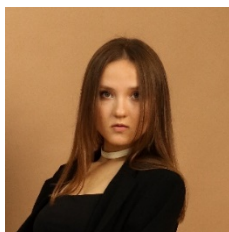
*Department of Information and Computer Systems Design  
Faculty of Computer Engineering  
Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: dmitriytavlyuy@gmail.com*

**Abstract.** The article discusses approaches to the implementation of an advertising company on the Internet, describes effective solutions at each stage, highlights the most optimal methods for achieving the goals of an advertising company. It is shown that it is advisable to use Web Analytics to measure and analyze the effectiveness of marketing campaigns.

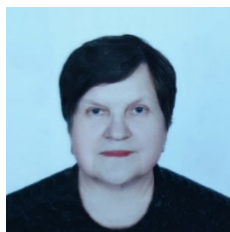
**Keywords:** Internet marketing, target advertising audience, competitor analysis, Web analytics.

УДК 621.396.96

## АВТОМАТИЗИРОВАННАЯ СИСТЕМА ИДЕНТИФИКАЦИИ ОБЪЕКТОВ ПО РАДИОЛОКАЦИОННЫМ И КАРТОГРАФИЧЕСКИМ ДАННЫМ



**В.С. Вашкевич**  
Студент Факультета  
Информационных технологий  
и управления БГУИР  
lera.vashkev18@gmail.com



**Л.Ф. Васковская**  
ОАО «АГАТ-системы  
управления» -  
управляющая компания  
холдинга  
«Геоинформационные  
системы управления»,  
ведущий системный  
аналитик  
lidavask@yandex.by



**А.В. Гордиевич**  
ОАО «АГАТ-системы  
управления» - управляющая  
компания холдинга  
«Геоинформационные  
системы управления»,  
инженер-программист  
alinagordievich78@gmail.com

### **В.С. Вашкевич**

Студентка 4 курса дневной формы обучения Белорусского Государственного Университета Информатики и Радиоэлектроники. Область научных интересов связана с разработкой алгоритмов и программного обеспечения для автоматизированных систем сбора данных и идентификации объектов.

### **Л.Ф. Васковская**

Окончила Минский радиотехнический институт. Область научных интересов связана с разработкой автоматизированных систем сбора и обработки информации о воздушной, наземной и надводной обстановках.

### **А.В. Гордиевич**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой методов и алгоритмов автоматизированной обработки информации для определения параметров случайных процессов.

**Аннотация.** Выполнена разработка алгоритма обработки радиолокационных данных об объектах воздушной, наземной и надводной обстановке и их объединения с картографической информацией для принятия решения об отнесении объектов к следующим классам: воздушный объект или наземный (надводный).

Рассмотрено применение статистических методов для анализа радиолокационных параметров местоположения и перемещения объектов, предложен алгоритм расчета интегральной оценки признаков для определения класса объекта по радиолокационной информации, разработан метод объединения интегральной оценки с картографическими данными для повышения уровня достоверности идентификации.

**Ключевые слова:** радиолокационные данные, картографические объекты, объекты инфраструктуры, интегральный рейтинг, тренд, среднеквадратические ошибки.

### **Введение.**

При радиолокационном наблюдении за воздушным пространством имеется проблема выделения таких объектов как обнаруживаемые наземные и (или) надводные цели [1]. К ним относятся, например, автомобильные средства, речные и морские суда, Эти цели должны быть идентифицированы и отделены от воздушных целей. В ходе выполненной работы предложен метод разделения наземных (надводных) и воздушных целей на основе анализа получаемых радиолокационных данных и картографических сведений о местности в части природного ландшафта, инфраструктуры, логистики. В основу метода положена



статистическая оценка параметров местоположения и движения обнаруживаемых радиолокационно объектов в процессе получения данных и принятие решения на основе интегрального рейтинга, рассчитываемого по совокупности оценок параметров и учета картографических сведений.

#### **Актуальность.**

В настоящее время отделение наземных (наводных) объектов от воздушных при наблюдении за окружающей обстановкой маловысотными РЛС является актуальной задачей в связи с расширением области применения радиолокационных средств, различных видов наземного и надводного транспорта, увеличением интенсивности и расширения разнообразия всех видов транспорта. Эти факторы требуют применения необходимых средства для различения каждого вида транспорта, специфического контроля за каждым из них, управления, обеспечения безопасности движения и принятия при необходимости адекватных мер воздействия.

Решение этих задач требует разработки автоматизированной системы [2], способной с помощью аппаратных и программных средств принимать решение на основе восприятия и анализа окружающей обстановки. Такая система должна обрабатывать информацию об обстановке от совокупности датчиков в реальном времени и своевременно представлять ее пользователю в соответствии с областью выполняемых им функций.

#### **Описание алгоритма идентификации по радиолокационным данным о координатах и скорости объекта.**

Для принятия решения о принадлежности объекта типу «воздушный» или «наземный (надводный)» по радиолокационным данным разработан алгоритм обработки параметров, состоящий из следующих действий.

1. Осуществляется сбор отметок от РЛС по каждой трассе наблюдаемого объекта за ряд обзоров. Набор отметок по отдельной трассе, собранный за ряд последовательных обзоров, представляет собой скользящее окно с фиксированным количеством отметок  $k$ . Это означает, что при каждом получении новой отметки трассы она заносится в окно, а самая старая по времени отметка удаляется из окна. Информация по каждой отметке трассы включает номер трассы, координаты  $X$ ,  $Y$ ,  $H$  и время их измерения. При приходе каждой отметки для расчетов используются отметки трассы на интервале получения отметок от  $i = n - k$  до  $i = n$ , где  $n$  – номер последней пришедшей отметки,  $k$  – размер скользящего окна.

2. Выполняются следующие операции по отметкам трассы, находящимся в скользящем окне:

– строится линия тренда трассы по высотам  $H_i$  отметок трассы, полученным в моменты времени  $i$ . Тренд рассчитывается в системе координат  $H0t$ : где  $H$  – вертикальная ось измерения высоты,  $t$  – горизонтальная ось измерения текущего времени,  $0$  – начало отсчета  $H$  и  $t$ . Эта линия позволяет определить сглаженное значение высоты  $H_m(i)$  объекта на момент получения последней текущей отметки по значениям высот отметок трассы в скользящем окне. Величина  $H_m(i)$  также называется скользящим средним.

Для определения  $H_m(i)$  в качестве линии тренда выбрана линия первого порядка, которая определяется по уравнению:

$$H_m(i) = at_i + b, \quad (1)$$

где  $a$  – тангенс угла наклона линии тренда по отношению к оси  $t_0$ ,

$b$  – координата точки пересечения линии тренда с осью  $H_0$ ;

$t_i$  – время получения отметки.

Для расчета коэффициентов  $a$  и  $b$  использован метод наименьших квадратов [3], который реализуется с помощью системы уравнений (2):



$$\begin{cases} a \sum_{i=n-k}^n (t_i) + b \cdot k = \sum_{i=n-k}^n (H_i) \\ a \sum_{i=n-k}^n (t_i^2) + b \sum_{i=n-k}^n (t_i) = \sum_{i=n-k}^n (H_i \cdot t_i) \end{cases} \quad (2)$$

где  $H_i$  – высоты отметок трассы.

Рассчитанное значение  $H_m(i)$  сравнивается с порогом  $H_{m(\text{пороговое})}$ , по результатам чего формируется бинарное значение коэффициента  $K_H$ :

- если  $H_m(i) \leq H_{m(\text{пороговое})}$ , то  $K_H = 1$ ,
- если  $H_m(i) > H_{m(\text{пороговое})}$ , то  $K_H = 0$ .

Значение  $K_H$  учитывается в расчете интегральной оценки;

– рассчитывается среднеквадратическая ошибка  $\sigma_h$  определения высоты по линии тренда с использованием отметок в скользящем окне на момент времени  $t$  получения последней  $i = n$  отметки по формуле:

$$\sigma_h = \sqrt{\frac{\sum_{i=n-k}^n (H_i - H_m(i))^2}{k}} \quad , \quad (3)$$

где  $k$  – количество отметок трассы в скользящем окне.

Рассчитанное значение  $\sigma_h$  сравнивается с порогом  $\sigma_{h(\text{пороговое})}$ , по результатам чего формируется бинарное значение коэффициента  $K_{\sigma_h}$ :

- если  $\sigma_h \leq \sigma_{h(\text{пороговое})}$ , то  $K_{\sigma_h} = 1$ ,
- если  $\sigma_h > \sigma_{h(\text{пороговое})}$ , то  $K_{\sigma_h} = 0$ .

Значение  $K_{\sigma_h}$  учитывается в расчете интегральной оценки;

– по линии тренда, рассчитанной по данным скользящего окна, определяется наличие или отсутствие набора или снижения высоты объекта на текущий момент  $t$  времени. Для этого на момент времени  $t$  вычисляется угол наклона  $\gamma$  линии тренда к оси  $t_0$  (в направлении оси от более раннего значения времени к более позднему) в системе координат  $H(t)$  по формуле

$$\pm\gamma = \arctg(a), \quad (4)$$

где  $a$  – коэффициент, который определяется при решении системы уравнений (2).

Положительное значение угла наклона ( $+\gamma$ ) соответствует набору высоты объектом, как показано на рисунке 1 в системе координат  $H(t)$ . Отрицательное значение угла наклона ( $-\gamma$ ) соответствует снижению высоты объекта, как показано на рисунке 2 также в системе координат  $H(t)$ .

Вычисленное значение угла наклона  $\pm\gamma$  сравнивается с порогами ( $+\gamma_{(\text{пороговое})}$ ) и ( $-\gamma_{(\text{пороговое})}$ ), по результатом чего формируется бинарное значение коэффициента  $K_\gamma$ :

- если  $(-\gamma_{(\text{пороговое})}) \leq (\pm\gamma) < (+\gamma_{(\text{пороговое})})$ , то  $K_\gamma = 1$ ,
- если  $(-\gamma_{(\text{пороговое})}) > (\pm\gamma) \geq (+\gamma_{(\text{пороговое})})$ , то  $K_\gamma = 0$ .

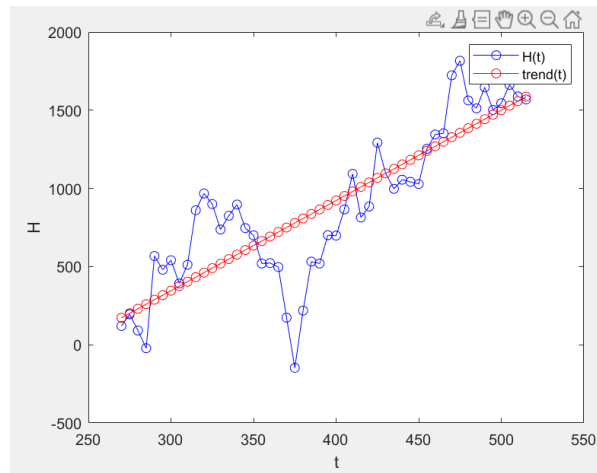


Рисунок 1. Линия тренда высоты  $H(t)$  для объекта, набирающего высоту

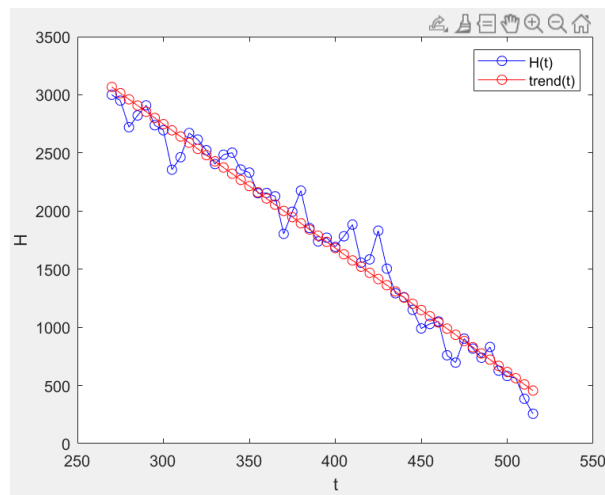


Рисунок 2. Линия тренда высоты  $H(t)$  для объекта, снижающего высоту

Значение  $K_\gamma$  учитывается в расчете интегральной оценки;

– рассчитывается среднеквадратическая ошибка  $\sigma_\gamma$  расчета угла наклона линии тренда по формуле:

$$\sigma_\gamma = \sqrt{\frac{\sum_{i=n-p}^n (\gamma_i - \gamma_m)^2}{p}}, \quad (5)$$

где  $\gamma_m$  – математического ожидание (среднее значение) угла наклона линии тренда высоты на участке трассы, представленном отметками в скользящем окне, рассчитывается по формуле:

$$\gamma_m = \frac{\sum_{i=n-p}^n \gamma_i}{p}, \quad (6)$$

где  $p$  – количество углов наклона тренда, учитываемых при расчете, учет которых начат от момента наличия по трассе не менее 4 отметок:  $p = k - 3$ .

Рассчитанное значение среднеквадратической ошибки  $\sigma_\gamma$  сравнивается с порогом  $\sigma_{\gamma(\text{пороговое})}$ , по результатам чего формируется бинарное значение коэффициента  $K\sigma_\gamma$ :

- если  $\sigma_\gamma \leq \sigma_{\gamma(\text{пороговое})}$ , то  $K\sigma_\gamma = 1$ ,
- если  $\sigma_\gamma > \sigma_{\gamma(\text{пороговое})}$ , то  $K\sigma_\gamma = 0$ .

Значение  $K\sigma_\gamma$  учитывается в расчете интегральной оценки;

– по координатам отметок трассы, накопленным в скользящем окне, рассчитывается средняя скорость  $V_m(i)$  (модуль скорости) движения объекта на момент получения последней  $i = n$  отметки в момент времени  $t$ . Расчет  $V_m(i)$  производится по составляющим скорости  $V_x(i)$ ,  $V_y(i)$  и модулям  $V(i)$  скорости каждой отметки с учетом всех отметок скользящего окна. Каждое значение  $V_x(i)$ ,  $V_y(i)$  рассчитывается по полученным от РЛС координатам  $XU$  соседних отметок трассы. Для расчета среднего значения  $V_m(i)$  используются следующие формулы:

$$V(i) = \sqrt{V_{x(i)}^2 + V_{y(i)}^2}, \quad (7)$$

$$V_x(i) = \frac{x_i - x_{i-1}}{t_i - t_{i-1}}, \quad (8)$$

$$V_y(i) = \frac{y_i - y_{i-1}}{t_i - t_{i-1}}, \quad (9)$$

$$V_m(i) = \frac{\sum_{j=n-l}^l V_j}{l}. \quad (10)$$

Рассчитанное значение  $V_m(i)$  сравнивается с порогом  $V_{m(\text{пороговое})}$ , по результатам чего формируется бинарное значение коэффициента  $K_v$ :

- если  $V_m(i) \leq V_{m(\text{пороговое})}$ , то  $K_v = 1$ ,
- если  $V_m(i) > V_{m(\text{пороговое})}$ , то  $K_v = 0$ .

Значение  $K_v$  учитывается в расчете интегральной оценки.

3. На основе полученных в пункте 2 значений коэффициентов  $K_H$ ,  $K\sigma_h$ ,  $K_\gamma$ ,  $K\sigma_\gamma$ ,  $K_v$  формируется **показатель идентификации  $S_t$  на момент времени  $t$** . Расчет показателя идентификации  $S_t$  производится по формуле (11):

$$S_t = K_H \cdot w_1 + K\sigma_h \cdot w_2 + K_\gamma \cdot w_3 + K\sigma_\gamma \cdot w_4 + K_v \cdot w_5, \quad (11)$$

где  $w_1, w_2, w_3, w_4, w_5$  – веса, присвоенные каждому коэффициенту  $K_H, K\sigma_h, K_\gamma, K\sigma_\gamma, K_v$ .

Весовые коэффициенты являются нормированными таким образом, что:

$$\sum_{i=1}^5 w_i = 1. \quad (12)$$

Значения весовых коэффициентов являются настраиваемыми величинами и их конкретные значения определяются в процессе моделирования.

4. Далее производятся вышеописанные расчеты значений показателей идентификации  $S$  для одного и того же объекта в нескольких обзорах РЛС. Таким образом, чтобы, например, к моменту времени  $t$  имелись значения показателей идентификации  $S_t, S_{t-1}, S_{t-2}, \dots, S_{t-p}$  за  $p$  предыдущих обзоров РЛС. По этим данным определяется интегральный показатель идентификации  $S$  по формуле:

$$S = S_t + S_{t-1} + S_{t-2} + \dots + S_{t-p}. \quad (13)$$

Полученное вычисленное значение интегрального показателя  $S$  сравнивается с порогом  $S_{(пороговое)}$  для определения бинарного значения  $U_t$ :

- если  $S \geq S_{(пороговое)}$ , то принимается решение  $U_t = 1$ ,
- если  $S < S_{(пороговое)}$ , то принимается решение  $U_t = 0$ .

В последующих  $m$  ( $m \geq 3$ ) обзорах производятся расчеты величин  $S$ , их сравнение с порогом  $S_{(пороговое)}$ , формирование бинарных значений  $U_{t+1}, \dots, U_{t+m}$ .

Решение об отнесении объекта к определенному классу – воздушный или наземный (надводный) производим по критерию « $k$  из  $m$ », где  $k$  – количество обзоров, в которых для интегрального показателя идентификации  $S$  выполнялись условия  $U_t = 1, U_{t+1} = 1, \dots, U_{t+m} = 1$ , (при этом всегда должно быть установлен критерий  $k \leq m$ ).

При выполнении критерия объект будет отнесен к классу «наземный (надводный)». При невыполнении критерия объект будет отнесен к классу «воздушный».

Если на момент времени проверки критерия « $k$  из  $m$ » для объекта выполняется условие:

$$\sigma_h > 2 \sigma_{h(пороговое)},$$

где  $\sigma_h$  определяется по формуле (3), то независимо от состояния выполнения (невыполнения) критерия « $k$  из  $m$ » для интегрального показателя идентификации принимается решение, что класс объекта не определен.

#### **Описание алгоритма идентификации по картографическим данным.**

В ряде случаев в зависимости от совокупности комплекса воздействующих внешних условий отделить воздушные объекты от наземных (надводных) по радиолокационным данным с требуемой степенью достоверности не удастся.

Для решения задачи в таких ситуациях положительный вклад может внести знание обстановки, которое предоставляет карта позиционно района, в котором работает РЛС.

Для использования карты как дополнительного источника информации об имеющихся над (на) территорией объектов для распознавания на карте позиционного района необходимо выделить объекты, представляющие интерес для разделения воздушных и наземных (надводных) объектов. Наличие этих картографических объектов поможет уточнить отнесение идентифицированных по радиолокационным данным объектов к обозначенным классам.

Алгоритм идентификации с учетом картографическим данным состоит в объединении результатов идентификации, полученных по радиолокационным данным, с картографической информацией о специальных картографических объектах.

Интерес для уточнения класса объекта представляют следующие картографические объекты:

- объекты инфраструктуры (площадные и линейные). К ним относятся автомобильные дороги, судоходные водоемы, строения, железнодорожные пути, аэропорты и аэродромы и т.п.;
- взлетно-посадочные полосы аэродромов;
- коридоры пролетов в воздушном пространстве государства;
- зоны, разрешенные и запрещенные для полетов авиации;
- рельеф местности.

В процессе использования картографической информации для ее объединения с информацией о распознанных по радиолокационной информации объектах осуществляются следующие этапы:

1. Считывание с карты координат местоположения картографических объектов – линейных площадных, которые указаны выше, на обозначенной территории.

2. Построение вокруг линейных и площадных объектов расширенных областей для последующего определения нахождения распознаваемых по радиолокации объектов в этих областях.

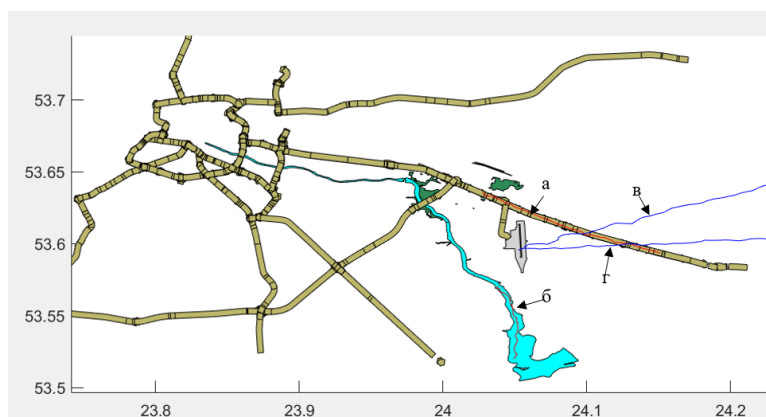
3. Для каждого распознанного по локационным данным объекта проверяется его нахождение во введенных ограничительных областях имеющих картографических объектов. Определяется его нахождение в ограничительных пределах, отведенных раздельно для наземных, надводных или воздушных объектов. Так, например, по взлетно-посадочной полосе преимущественно могут перемещаться воздушные объекты, над водными поверхностями - более вероятно перемещение надводных объектов, вдоль автомобильных дорог – в основном наземный автомобильный транспорт.

По результатам проверки определяется в область какого картографического объекта попадает исследуемый объект. Проверка производится несколько раз и конечный результат определяется по критерию « $k$  из  $m$ ».

Результат классификации, полученный по карте, объединяется с результатом классификации по интегральному радиолокационному показателю. Картографический анализ помогает более точно провести классификацию объектов. Особенно важна роль карты при решении спорных ситуаций об отнесении объекта к воздушному или наземному (надводному).

#### **Моделирование использования картографической информации при идентификации объектов.**

Для примера использована карта реального позиционного района в Беларуси.



а – наземный объект, б – надводный, в и г – воздушные объекты

Рисунок 3. Карта позиционного района

На карте изображены следующие объекты: река, автомобильные дороги, аэропорт.

С помощью имитатора для моделирования были заданы: два воздушных объекта, взлетающих из аэропорта (отмечены как **в** и **г** на рисунке 3) и имеющих среднюю скорость 200 км/час, наземный объект (отмечен как **а** на рисунке 3), движущийся по трассе со скоростью 180 км/час, надводный объект (отмечен как **б** на рисунке 3), движущийся вдоль русла реки со скоростью 100 км/час.

Картографические объекты были извлечены из открытого интернет-ресурса *OpenStreetMap* в векторном формате (*ESRI Shape File*) с использованием *QGIS Desktop* – свободной кроссплатформенной геоинформационной системы, предназначенной для редактирования, визуализации, анализа и публикации геопространственной информации.

Дороги были представлены как линейные объекты, заданные в виде отрезков линий, аэродром и река - как площадные объекты.

На аэродроме была задана точка стояния РЛС, которая принята за начало (0) системы отсчета координат  $XU$ , в которой происходит наблюдение за трассами объектов. Необходимые пересчёты из географической системы координат в прямоугольную и обратно осуществлялись с использованием встроенных функций в *MATLAB*.

Признаками, используемыми на данном этапе идентификации, являются области, заданные последовательностью точек на карте с координатами широты и долготы. Решение задачи сводится к проверке на попадание идентифицируемого объекта в какую-либо из заданных областей на карте. Вокруг линии дороги были заданы прямоугольные области шириной 500 метров (с учетом ошибки в определении координат объекта). Вокруг площадных объектов также были заданы расширенные области для учета ошибок определения координат объектов.

Данные о попадании или непопадании объекта в каждый вид областей (дороги, аэродром, река) получаем для каждой пришедшей отметки: если пришедшая в момент времени  $T_i$  отметка с координатами  $X_i, Y_i, H_i$  попала в область дороги, то оценка картографических признаков  $i$ -той отметки по каждой из областей будет выглядеть, как показано в таблице 1.

Таблица 1. Оценка каждой отметки на попадание в область

Номер отметки	Признак		
	Дорога	Аэродром	Река
$i$	1	0	0

В итоге каждой пришедшей отметке может соответствовать только один объект карты.

Эти оценки объединяются вместе при наблюдении за трассой каждого объекта следующим образом. Каждому из признаков (областей) присваивается номер: дорога – 1, река – 2, аэродром – 3, и если не в одну из этих областей отметка не попала, признак принимает значение 0. Информация о попадании 5 отметок трассы воздушного объекта может выглядеть как указано в таблице 2.

Таблица 2. Вид получения картографической информации об объекте

Номер отметки	Номер области, в которую попала отметка
1	3
2	3
3	0
4	0
5	1

На следующем шаге результат, полученный по расчетным признакам (1 – наземный объект, 0 – воздушный, 10 – результат неопределен), сравнивается с картографической информацией. Картографическая информация может улучшать результат идентификации, полученный по радиолокационным данным, но не ухудшать его. Поэтому проверять на попадание в выделенные области будем объекты, принятые за наземные, и объекты, для которых класс не определен.

Пусть пришла отметка объекта, результат идентификации которой равен 1 (наземный или надводный объект). Проверяем, в какую область попала отметка: если попала в 1 или 2 (дорога или река) – результат идентификации подтверждается и равен 1 (наземный или надводный объект). Если отметка попала в области 3 и 0 (аэродром и территория вне

заданных областей), проверяем 5 последних отметок на попадание по критерию « $k$  из  $m$ », где  $k \geq 3$ , а  $m = 5$ , то есть для подтверждения, что объект наземный, необходимо чтобы как минимум 3 из 5 последних отметок объекта попали в область дороги или реки.

#### **Заключение.**

В статье рассмотрена возможность создания автоматизированной системы идентификации объектов совместно по радиолокационным и картографическим данным. Система предназначена для определения таких классов объектов как воздушные, наземные и (или) надводные. Источником радиолокационных данных является РЛС, осуществляющая обнаружение и сопровождение трасс всех наблюдаемых объектов. В составе картографических данных использованы сведения о координатах местонахождения объектов, которые предназначены для преимущественного использования специализированными видами транспорта, такими как авиация, автомобильный транспорт, надводные суда. Для распознавания применены также характеристики природного ландшафта местности. В основу классификации по радиолокационным данным положено определение интегрального показателя, использующего такие параметры объектов как высота и скорость, которые накапливаются по радиолокационным трассовым отметкам в скользящем окне. Объединение интегрального показателя с картографической информацией обеспечило повышение достоверности распознавания. Для подтверждения эффективности разработанного алгоритма идентификации выполнено моделирование в среде MATLAB.

Полученные результаты предоставляют возможность для выполнения автоматизированной идентификации объектов в объеме указанных классов.

#### **Список литературы**

- [1] INTRODUCTION TO RADAR SYSTEMS International edition 2001 third edition / Merrill I. Skolnik / McGRAW-HILL INTERNATIONAL EDITION / Electrical Engineering Series, 2001
- [2] Фор, А. Восприятие и распознавание образов / Пер. С фр. А. В. Серединского; под ред. Г. П. Каатыса. – М.: Машиностроение, 1989. – 272 с.
- [3] Бронштейн, И. Н. Справочник по математике для инженеров и учащихся втузов. Издание третье, переработанное / И. Н. Бронштейн, К. А. Семендяев. – М.: ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО ТЕХНИКО-ТЕОРЕТИЧЕСКОЙ ЛИТЕРАТУРЫ 1953. – 608 с.

## **AUTOMATED OBJECTS IDENTIFICATION USING RADAR AND MAP DATA**

***V. S. Vashkevich***

*BSUIR, Faculty of Information  
Technology and Control student*

***L. F. Vaskovskaya***

*Open joint-stock company  
«AGAT – Control Systems –  
Managing Company of  
Geoinformation Control  
Systems Holding», leading  
system analyst*

***A. V. Gordievich***

*Open joint-stock company  
«AGAT – Control Systems –  
Managing Company of  
Geoinformation Control  
Systems Holding», engineer-  
programmer*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: lera.vashkev18@gmail.com*

**Abstract.** The algorithm of aerial and ground (overwater) objects' radar data processing and joining it with map data for making a decision to relate the object to the one of the defined classes: ground (overwater) object and aerial object.

Statistics methods in object's coordinate and movement data analysis are shown, class determination algorithm for attributes integral estimation, using radar data is proposed, method of integral estimation and map data combination is developed for increasing the level of the identification accuracy.

**Keywords:** radar data, map objects, infrastructure objects, integral rating, trend, mean square errors.



УДК 004.045

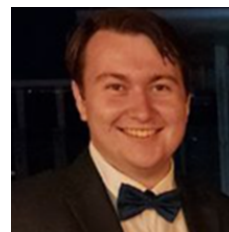
## ОЦЕНКА ДОВЕРИЯ К РАЗЛИЧНЫМ ИНФОРМАЦИОННЫМ ИСТОЧНИКАМ И МЕТОДЫ РАБОТЫ С НИМИ В ЭПОХУ BIG DATA



**Н.А. Калабин**  
Студент 1 курса  
специальности «Информатика  
и технологии  
программирования»  
5203330@gmail.com



**А.О. Чмутов**  
Студент 1 курса  
специальности «Информатика  
и технологии  
программирования»



**В.Д. Владыцев**  
Ассистент кафедры  
информатики, инженер-  
программист ОИАСУ ЦИИР  
БГУИР  
v.vladymtsev@bsuir.by

**Н.А. Калабин**  
Студент 1-го курса БГУИР ФКСИС, по специальности ИИТП

**А.О. Чмутов**  
Студент 1-го курса БГУИР ФКСИС, по специальности ИИТП

**В.Д. Владыцев**  
Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР

**Аннотация.** Получать информацию стало проще, но многие источники теперь не отличаются качеством, что означает необходимость проведения надежных исследований. Дополнительная проблема заключается в том, что любой человек может распространять дезинформацию, что усложняет задачу отличить правду от лжи. Анализ больших данных может помочь в определении эффективных источников и методов получения информации, что впоследствии может привести к разработке более эффективных и надежных стратегий использования этих ресурсов.

**Ключевые слова:** Big Data, Большие Данные, Источники информации, Дезинформация, Доверие к источникам информации, Влияние.

### Введение.

С приходом века высоких технологий получить доступ к информации стало как никогда просто, информация буквально окружает нас. Источники информации представлены в огромном количестве форм, таких как люди, документы, книги, веб-сайты, видео и различные другие носители знаний. Каждый из них предоставляет данные о конкретной теме или предмете. Сложно переоценить ценность информации. В умелых руках информация используется как инструмент войны: с её помощью свергают режимы, уничтожают человеческие судьбы, лоббируют нужные решения.

### Дезинформация и роль аналитики больших данных в борьбе с ней.

Каналы средств массовой информации принято делить на две группы: печатные СМИ (газеты, журналы, книги и другие задокументированные формы) и электронные, которые, в свою очередь, включают как «традиционные» вещательные (радио и телевидение), так и «новые» медиа (социальные сети) [1], а также мы выделяем нейронные сети.

Традиционные источники информации, хоть и потеряли сегодня свою былую популярность [2], по-прежнему оказывают значительное влияние на восприятие, мнения и убеждения людей [1]. Однако, они имеют несколько проблем:

Традиционные источники часто основаны на доминирующих культурных нормах и ценностях, которые могут исключать или маргинализировать голоса, не соответствующие

этим нормам. Более того, традиционные источники информации подвержены цензуре, особенно в авторитарных режимах, что может ограничить доступ к информации, которая бросает вызов статусу-кво.

Кроме того, интерпретация традиционных источников информации часто зависит от индивидуальных предубеждений и мировоззрения читателя. Такая субъективность может привести к неправильному толкованию или искажению информации, увековечиванию дезинформации и укреплению общественных стереотипов.

Для контекстуализации этих вопросов стоит рассмотреть пример исторического текста, такого как "Одиссея" Гомера. Хотя "Одиссея" считается классическим произведением литературы, она также является продуктом своего времени, отражая доминирующие культурные нормы и ценности Древней Греции. Однако ее темы героизма и путешествия к самопознанию продолжают находить отклик у современных читателей. Поэтому традиционные источники информации следует анализировать критически, осознавая их контекст и ограниченность их перспектив.

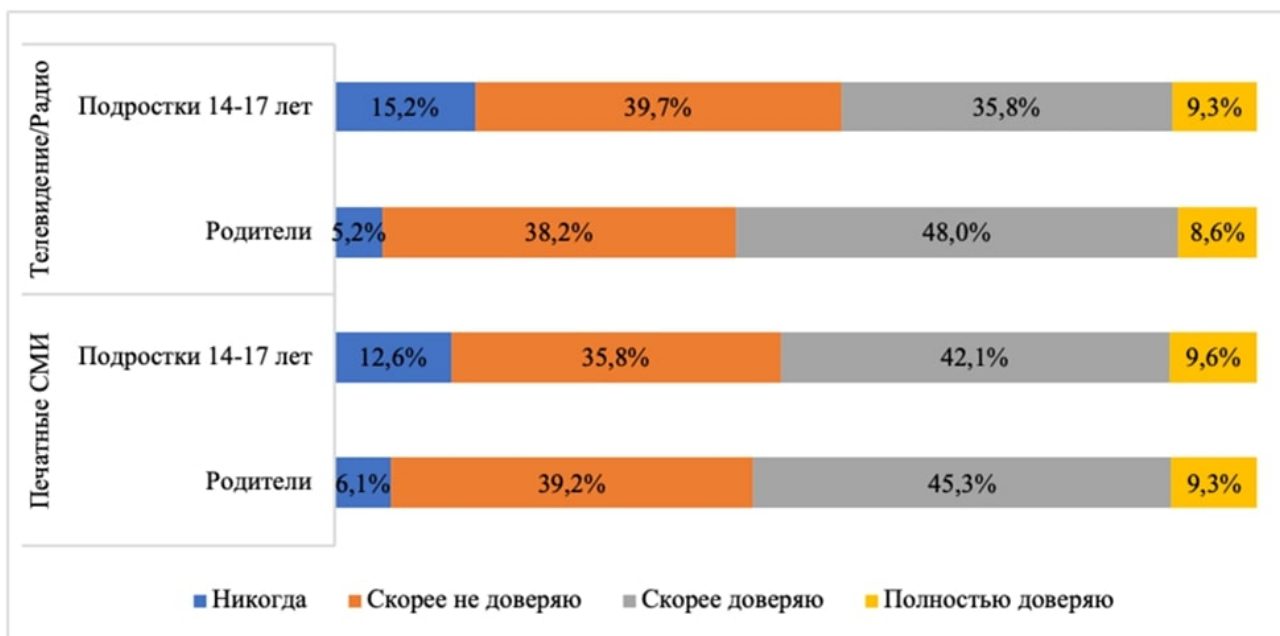


Рисунок 1. Доверие к информации из онлайн-источников у подростков и родителей, %  
[1, с14]

Новые медиа сегодня преобладают над традиционными [2], они охватывают гигантскую аудиторию в реальном времени и делают информацию общедоступной. Доступность информации снизила ее качество, что привело к необходимости проводить собственные мини-исследования для получения достоверных сведений. Мы должны точно определить первоисточник и отличить надежные источники от ненадежных, чтобы получить достоверную информацию. Это занимает время и отнимает много сил, особенно у людей, которые родились до появления интернета. При этом нет гарантий, что вы найдете верную информацию. Эта проблема усугубляется тем, что любой человек, независимо от аккредитации или образования, теперь может выступать в качестве источника информации, т. к. увеличивается роль социальных сетей: в американском исследовании каждый второй взрослый называл их для себя постоянным источником новостной информации [3]. К сожалению, наказания за фабрикацию и распространение фальшивых историй незначительны по сравнению с огромным негативным воздействием, которое может

оказать дезинформация. В 2013 г. Всемирный экономический форум включил массовую цифровую дезинформацию в число наиболее серьезных глобальных рисков [4].

В силу того, какую роль сегодня играют социальные сети, они привлекают к себе много внимания, исследователи и практики стремятся понять динамику онлайн коммуникации, однако сталкиваются со множеством проблем с анализом данных:

Одной из проблем при анализе информации является качество самих данных. Платформы ежедневно генерируют большое количество информации, которая зачастую не релевантна, неточна или нерепрезентативна [5].

Однако данные социальных сетей не отражают мнение или поведение пользователей. Влияние оказывают социальные нормы, давление со стороны или самопрезентация. Необходимо применять строгие методы и техники для фильтрации, проверки и подтверждения данных [6].

Еще одной проблемой является сбор данных. Социальные сети часто накладывают ограничения на доступ к информации. Например, ограничивают количество запросов, временные рамки или объем, которые можно получить через API [5].

Кроме того, платформы социальных сетей могут менять свою политику, алгоритмы или функции без предварительного уведомления, что оказывает влияние на последовательность и сопоставимость.

Платформы не предоставляют достаточной информации о местоположении или структуре сети. Поэтому исследователи и аналитики нередко должны знать об ограничениях и проблемах, приводящих к поиску альтернативных источников.

Третья проблема — интерпретация данных. Данные часто являются сложными и многомерными, включая различные типы и форматы контента. Они часто неоднозначны и зависят от контекста, что требует глубокого понимания лингвистических, культурных и ситуационных аспектов коммуникации. Например, значение хэштега или эмодзи может меняться в зависимости от пользователя или платформы. Необходимо применять соответствующие методы и инструменты для анализа данных таким образом, чтобы отобразить богатство и разнообразие, а также учесть их неопределенность и изменчивость.

Четвертая проблема анализа информации из социальных сетей — этика данных. Данные социальных сетей включают персональную информацию о личности, предпочтениях, мнениях и поведении пользователей, что поднимает такие этические вопросы, как неприкосновенность частной жизни, согласие, анонимность, конфиденциальность и право собственности. Более того, данные социальных сетей могут иметь последствия для репутации, безопасности, прав и благополучия пользователей. Поэтому исследователям и аналитикам необходимо следовать этическим принципам и рекомендациям, чтобы защитить достоинства заинтересованных сторон.

Анализ информации из социальных сетей является сложным процессом, требующим тщательного рассмотрения различных аспектов. Эти критерии можно рассматривать и как возможности для разработки новых и инновационных подходов, повышающих ценность и влияние аналитики социальных сетей.

Социологические исследования показали, что подростки в возрасте 14–17 лет имеют высокий уровень доверия к социальным сетям, как источникам информации, так каждый второй сообщает о среднем уровне доверия и 11% полностью доверяют ленте в социальных сетях [1, с.14]. Это чрезмерное доверие молодого поколения часто приводит к тому, что они верят ложной информации и делятся ею без предварительной проверки.

По сравнению с детьми, у родителей уровень доверия к онлайн-ресурсам ниже, а к традиционным медиа — выше, что подтверждается и в зарубежных исследованиях [7].

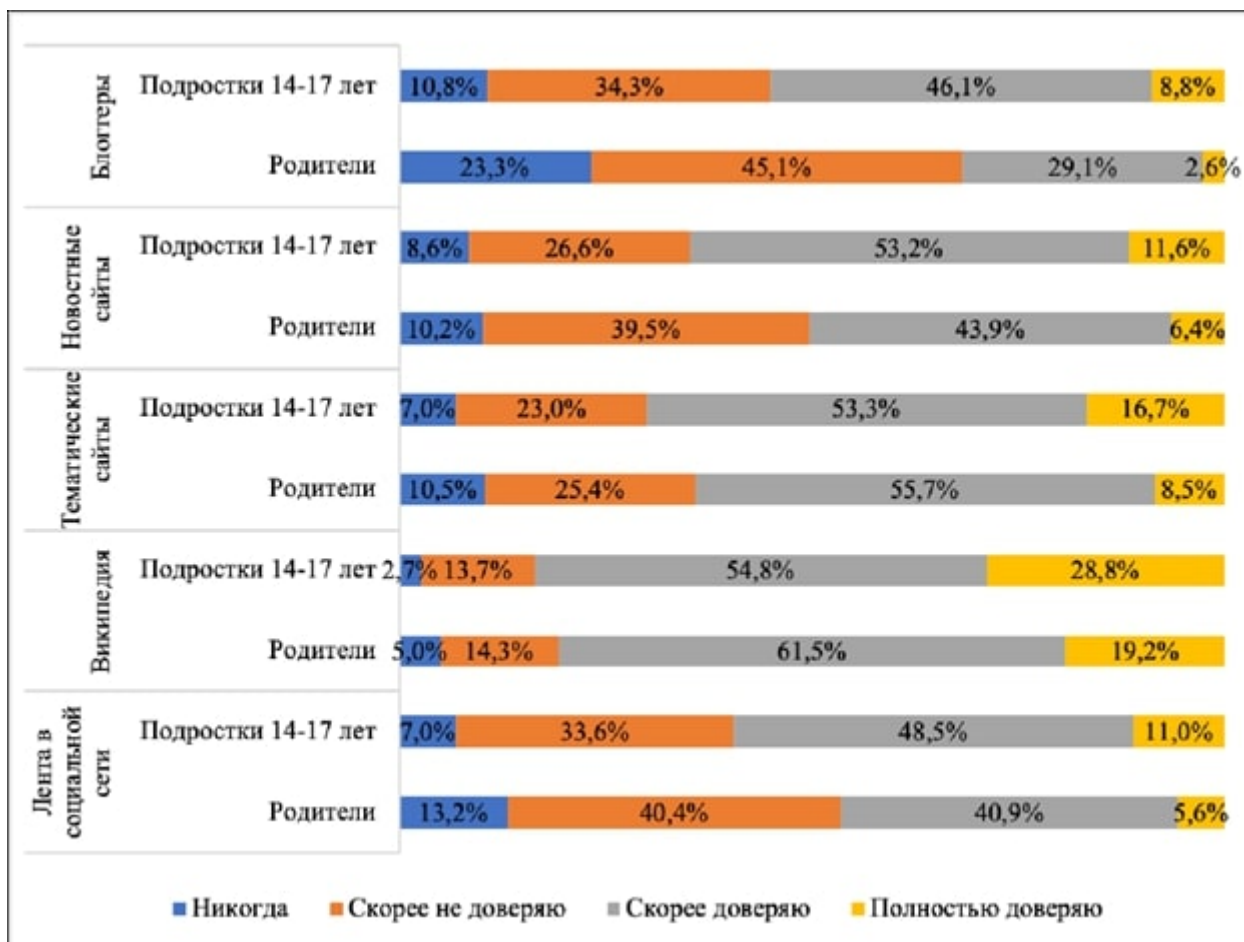


Рисунок 2. Доверие к информации из онлайн-источников у подростков и родителей, %.  
[1, с14]

Невозможно не учесть, как источник информации, появившиеся совсем недавно, нейронные сети, основанные на методе глубокого обучения, они позволяют человеку за считанные секунды получить конкретный ответ на свой вопрос. На данный момент, Chat GPT, в основе которого лежит мультимодальная большая языковая модель GPT-4 [8], является наиболее продвинутой нейронной сетью, она проходит собеседования на ведущие позиции в ИТ компании, сдаёт экзамены на уровне 10% лучших выпускников ВУЗов [9]. Кроме GPT-4 существует множество других языковых моделей, таких как BERT, T-NLG, Switch-C, ELMo и др. Но они все уступают GPT-4 на данный момент как по количеству параметров, так и по качеству и точности ответов.

При вопросе языковая модель анализирует текст, чтобы воспринять его и ответить соответствующим образом. Её понимание языка и контекста позволяет создать ответ, предсказывая идеальную последовательность слов, извлеченных из массы текстовых данных, которые она изучила. По сути, языковая модель генерирует ответ, предсказывая, каким будет следующее наиболее вероятное слово(а), основываясь на вводимом тексте и контексте разговора. Она также может генерировать несколько возможных ответов и выбирать наиболее подходящий на основе различных факторов, таких как релевантность, согласованность и грамматическая корректность.

Одной из основных проблем языковых моделей является то, что они очень сильно зависят от той информации, которую им предоставили в момент обучения, ведь именно на её основании строятся ответы [10]. Например, исследователи обнаружили, что BERT

ассоциирует фразы, в которых упоминаются инвалиды, с более негативными словами настроения, и что насилие с применением оружия, бездомность и наркомания чрезмерно представлены в текстах, обсуждающих психические заболевания [11]. Аналогично, модели типа GPT-3, обученные на не менее чем 570 ГБ данных, полученных в основном из Common Crawl 16 [12], могут генерировать предложения с высокими оценками токсичности, даже когда им предлагаются нетоксичные предложения [13]. А при изучении обучающих данных для GPT-217 обнаружили 272 тыс. документов с ненадежных новостных сайтов и 63 тыс. документов из запрещенных веток Реддита. Из-за всего вышеперечисленного ответы на некоторые вопросы могут быть некорректными или ошибочными.

Второй проблемой можно назвать контроль, хоть в этом направлении и предпринимаются серьезные усилия [9], но языковые модели всё ещё можно взломать и получить от них запрещенную информацию.

Третья проблема — это сложность интерпретации. Языковые модели могут быть сложны в интерпретации из-за их нелинейности и сложных внутренних структур, что затрудняет объяснение принятых решений и выводов.

GPT-4 и другие языковые модели на основе глубокого обучения — это невероятно мощные и сложные инструменты, которые отвечают на запросы пользователей, используя огромные объемы текстовых данных. Несмотря на их сложность, они не лишены ограничений. Например, GPT-3 может столкнуться с проблемами контроля, интерпретации и ошибками из-за зависимости от обучающих данных. Важно подходить к этим моделям с осторожностью, учитывая их ограничения и недостатки.

При углублении в научные исследования, посвященные доверию к различным источникам информации и эффективным механизмам преодоления трудностей стоит изучить применение аналитики больших данных.

Анализ огромных массивов данных позволяет выявить охват и эффективность источников информации в различных областях жизни, включая науку, бизнес, медицину, политику и многое другое — потенциальная золотая жила для понимания социальных тенденций. Обработка больших данных позволяет выявить модели потребления и предпочтения аудитории, основанные на различных переменных, таких как возраст, пол, образование или местоположение.

Существует несколько подходов к изучению использования источников информации в различных областях и выявлению наиболее эффективных методов для достижения конкретных целей. Таких как:

Исследование пользователей — это распространенный метод, используемый для сбора информации среди пользователей, которые ищут решения различных проблем в своей профессиональной или личной жизни. Чтобы получить эти данные, исследователи проводят ряд исследований, которые могут включать в себя опросы, интервью, фокус-группы и тестирование продуктов. Конечный результат исследования играет ключевую роль в определении наиболее эффективных источников информации в конкретном сценарии.

Анализ данных — это альтернативный метод, который предполагает глубокое изучение использования различных источников информации. Можно анализировать данные, связанные с использованием различных источников информации в рабочей среде, таких как журналы посещений веб-сайтов, статистику использования информационных систем и другие метрики. Анализ данных может помочь определить наиболее эффективные источники информации в конкретной ситуации.

Исследование лучших практик: одним из эффективных методов изучения источников информации является исследование передового опыта в различных областях деятельности. Изучая, как организации и компании эффективно используют свои информационные



ресурсы, вы можете получить ценные знания, которые можно применить в собственной работе.

Экспертное мнение: консультации с экспертами в вашей области могут дать ценные сведения об эффективных источниках информации для достижения ваших целей и задач. Делясь своим опытом и советами в конкретных ситуациях, эксперты могут предложить уникальные рекомендации по источникам информации, которые могут быть непредсказуемыми, но плодотворными для расширения ваших знаний и навыков.

Определение наиболее эффективных источников информации в конкретном сценарии имеет ключевое значение и может быть осуществлено с помощью различных методов. Для получения всестороннего анализа рекомендуется применять комбинацию различных подходов, имеющихся в вашем распоряжении. Исследование больших данных может дать представление о частоте и эффективности различных методов и практик, связанных с источниками информации, в различных секторах.

### **Заключение.**

С помощью анализа больших данных можно выявить тенденции в использовании источников информации и методов работы с ними, определить их эффективность в различных сферах бизнеса, что приведет к разработке более мощных и надежных методов работы с этими ресурсами.

### **Список литературы**

[1] Солдатова Г.У., Чигарькова С.В., Илюхина С.Н. Медиапотребление подростков и родителей: источники информации и доверие к ним // Психологические исследования. 2021. Т. 14, № 80. С. 4. URL: <http://psystudy.ru/index.php/num/2021v14n80/1955-soldatova80.html> с-14.

[2] Медиапотребление в России-2020, URL: <https://oohmag.ru/wp-content/uploads/2020/11/mediapotrebienie-v-rossii-2020.pdf>

[3] Shearer E., Mitchell A. News Use Across Social Media Platforms in 2020 // Pew Research Center, 2021. URL: <https://www.journalism.org/2021/01/12/news-use-across-social-media-platforms-in-2020/>

[4] Howell L. Digital wildfires in a hyperconnected world // WEF Report 2013. World Economic Forum. Eight editions. 2013. 25 p.

[5] Stieglitz Stefan. The Adoption of Social Media Analytics for Crisis Management — Challenges and Opportunities, 2018. URL: [https://www.researchgate.net/publication/325416290\\_The\\_Adoption\\_of\\_Social\\_Media\\_Analytics\\_for\\_Crisis\\_Management\\_-\\_Challenges\\_and\\_Opportunities](https://www.researchgate.net/publication/325416290_The_Adoption_of_Social_Media_Analytics_for_Crisis_Management_-_Challenges_and_Opportunities)

[6] Weiguo Fan, Michael D. Gordon. Unveiling the Power of Social Media Analytics, 2014. URL: [https://www.researchgate.net/publication/259148570\\_The\\_Power\\_of\\_Social\\_Media\\_Analytics](https://www.researchgate.net/publication/259148570_The_Power_of_Social_Media_Analytics)

[7] Edelman Trust Barometer. Global Report, 2019, 65 p.

[8] GPT-4 — Википедия (wikipedia.org), URL: <https://ru.wikipedia.org/wiki/GPT-4>

[9] GPT-4 Technical Report, url: <https://arxiv.org/abs/2303.08774>

[10] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency (pp. 610-623).

[11] Ben Hutchinson, Vinodkumar Prabhakaran, Emily Denton, Kellie Webster, Yu Zhong, and Stephen Denuyl. 2020. Social Biases in NLP Models as Barriers for Persons with Disabilities. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Online, 5491–5501. <https://doi.org/10.18653/v1/2020.acl-main.487>

[12] <https://commoncrawl.org/>

[13] Samuel Gehman, Suchin Gururangan, Maarten Sap, Yejin Choi, and Noah A. Smith. 2020. RealToxicityPrompts: Evaluating Neural Toxic Degeneration in Language Models. In Findings of the Association for Computational Linguistics: EMNLP 2020. Association for Computational Linguistics, Online, 3356–3369. <https://doi.org/10.18653/v1/2020.findings-emnlp.301>

[14] Pause Giant AI Experiments: An Open Letter, URL: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

## **ASSESSMENT OF TRUST IN VARIOUS INFORMATION SOURCES AND METHODS OF WORKING WITH THEM IN THE ERA OF BIG DATA**

**N.A. Kalabin**  
*BSUIR student,  
system programmer engineer*

**A.O. Chmutau**  
*BSUIR student,  
system programmer engineer*

**V.D. Vladymtsev**  
*Assistant of the Department of  
Computer Science, Software  
Engineer of DIACS CIIR BSUIR*

*Belarusian state university of informatics and radioelectronics, Republic of Belarus  
E-mail: 5203330@gmail.com, v.vladymtsev@bsuir.by*

**Annotation.** It has become easier to get information, but many sources are no longer of high quality, which means that reliable research is needed. An additional problem is that anyone can spread disinformation, which complicates the task of distinguishing truth from lies. Big data analysis can help in identifying effective sources and methods of obtaining information, which can subsequently lead to the development of more effective and reliable strategies for using these resources.

**Keywords:** Big Data, Information sources, Disinformation, Trust in information sources, Influence.

УДК 004.021:004.75

## АЛГОРИТМ НАХОЖДЕНИЯ ОПТИМАЛЬНОГО ПУТИ В ТРАНСПОРТНОЙ СЕТИ В УСЛОВИЯХ ИЗМЕНЯЕМЫХ ПАРАМЕТРОВ



**В.П. Корячко**  
Заведующий кафедрой  
систем  
автоматизированного  
проектирования  
вычислительных средств,  
доктор технических наук,  
профессор  
koryachko.v.p@rsreu.ru



**А.В. Бакулев**  
Доцент кафедры систем  
автоматизированного  
проектирования  
вычислительных средств,  
кандидат технических  
наук, доцент  
alex.bakulev@gmail.com



**М.А. Бакулева**  
Доцент кафедры систем  
автоматизированного  
проектирования  
вычислительных средств,  
кандидат технических наук,  
доцент  
marina.bakuleva@gmail.com

### **В.П. Корячко**

Заслуженный деятель науки и техники РФ. Автор более 470 научных и учебно-методических работ, в том числе шести учебников и четырех монографий. Область научных интересов: системы автоматизированного проектирования, интеллектуальные системы, мягкие вычисления.

### **А.В. Бакулев**

Доцент кафедры САПР ВС РГРТУ имени В.Ф. Уткина. Автор более 100 научных и учебно-методических работ. Область научных интересов: структуры и алгоритмы данных, лингвистическое обеспечение система автоматизированного проектирования, конкретная математика, структурное программирование.

### **М.А. Бакулева**

Доцент кафедры САПР ВС РГРТУ имени В.Ф. Уткина, автор более 70 научных и учебно-методических работ. Область научных интересов: системы искусственного интеллекта, нечеткая логика, машинное обучение.

**Аннотация.** В работе представлен алгоритм нахождения оптимального пути, с функционалом адаптивного решения задачи оптимизации при условии постоянно меняющихся внешних факторов. Управляющие параметры (факторы) могут иметь различную семантику и соответствующие метрики, качестве примера выбраны наиболее распространённые – время в пути и пропускная способность транспортной сети. В качестве математической модели, позволяющей адекватно отразить многогранности внешних факторов, используется тензорная модель. В качестве базового алгоритма оптимизации предлагается распространённый алгоритм Флойда – Уоршелла. Данный алгоритм модифицирован в соответствии с поставленными задачами исследования.

**Ключевые слова:** транспортная сеть, пропускная способность, тензорная модель, модели загрузки транспортной сети, графы, алгоритмы на графах, кратчайший путь, алгоритм Флойда - Уоршелла.

### **Введение.**

Транспортная инфраструктура – важнейших из элементов жизни городской среды, регионов и стран. В последнее время отмечается ускоряющиеся темпы роста количества личного и общественного транспорта, что приводит к таким проблемам как замедление трафика передвижений, нарушению целостности дорожного полотна, увеличению количества аварий на дорогах и проч. Очевидно, что научные методы решения задачи оптимизации трафика могут быть реализованы только за счет разработки методов и алгоритмов эффективной навигации транспортных средств. С точки зрения математической постановки задачи необходимо разработать модель, позволяющую учитывать вариативность управляющих факторов, их многомерность и динамику.



В данной работе приводится математическая модель отвечающая многомерной парадигме: это тензорное представление многомерной структуры, интегрирующая матричный способ моделирования транспортных сетей и многомерность управляющих параметров.

Научная новизна представленного исследования заключается в обеспечении возможности учета характеристик транспортной сети, сложно формализуемого вероятностного характера. В матрице смежности транспортной сети веса рёбер представляют собой массив характеристик транспортной сети. Таким образом, задача исследования – это нахождение такого пути между двумя вершинами ориентированного графа, который учитывает критерий минимальности в интегрированном массиве веса каждой дуги.

В качестве данных подобных массивов могут выступать средние скорости движения, задержки, объёмы транспортировки грузов и т.д., то есть те величины, которые необходимо оптимизировать и которые обладают свойством аддитивности [1].

В качестве математической модели для транспортной сети обычно выбирается орграф  $G(X, V)$ , где  $X$  - множество вершин,  $V$  - множество дуг сети. Дуги ориентированного графа соответствуют дорогам, соединяющим перекрёстки, станции и т. д., то есть непрерывные отрезки пути. Все дуги сходятся в вершинах, которые представляют собой любые транспортные развязки. В большинстве случаев дороги направлены в обе стороны, поэтому на графе обычно они обозначаются направленной в обе стороны парой дуг (рисунок 1).

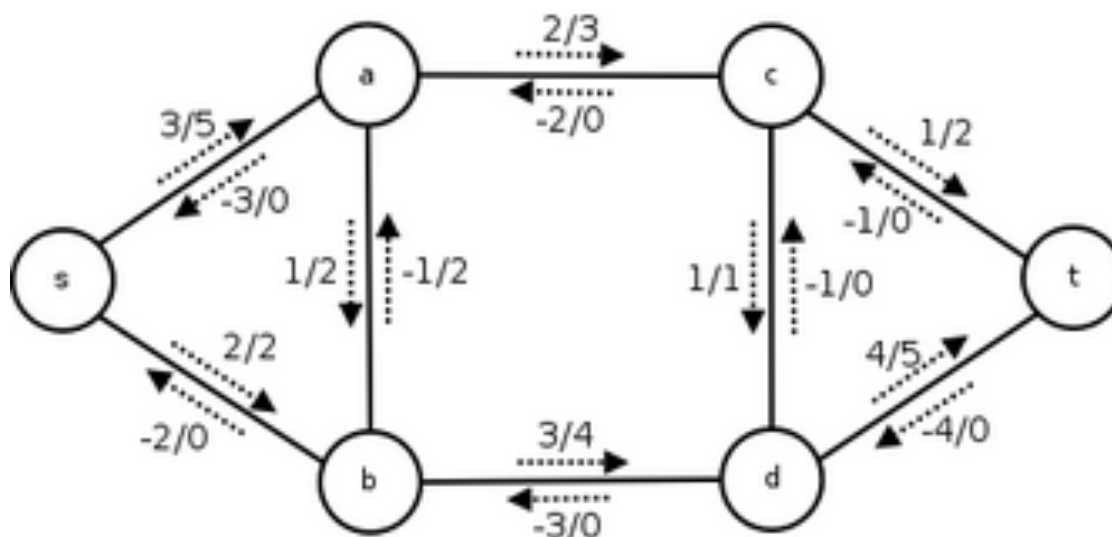


Рисунок 1. Пример графа абстрактной транспортной сети

В представленном графе рассматриваются и сопоставляются две оценки: например, длина пути и время задержки (простоя в пробке).

Математическим представлением данного графа будет матрица смежности следующей структуры:

$$G = \begin{pmatrix} & s & a & b & c & d & t \\ s & \infty & \|3 & 5\| & \|2 & 2\| & \infty & \infty & \infty \\ a & \|3 & 0\| & \infty & \|1 & 2\| & \|2 & 3\| & \infty & \infty \\ b & \|2 & 0\| & \|1 & 2\| & \infty & \infty & \|3 & 4\| & \infty \\ c & \infty & \|2 & 0\| & \infty & \infty & \|1 & 1\| & \|1 & 2\| \\ d & \infty & \infty & \|3 & 0\| & \|1 & 0\| & \infty & \|4 & 5\| \\ t & \infty & \infty & \infty & \|1 & 0\| & \|4 & 0\| & \infty \end{pmatrix}$$

Для описания данной матрицы предлагается использовать тензорную модель, адекватно отображающую многомерную структуру параметров транспортной сети [3]. На основе тензорной математической модели разработан алгоритм нахождения оптимального пути.

#### Актуальность.

На сегодняшний день известны подходы к решению задачи многокритериальной оптимизации маршрутов с теми или иными допущениями в зависимости от особенностей и приоритетных характеристик оптимизации. Наиболее распространенным является подход, основанный на расчёте весов графа, учитывающих все параметры оптимизации, с последующим нахождением оптимального пути, минимизирующего единственный параметр [2,3]. Аналогичные задачи нашли отражения в работах [4], где рассматривается модель управления дорожным потоком, интегрированная с системой ГИС. Особое значение в сетевом управлении имеет мониторинг характеристик транспортных потоков, дорожной сети, устройств управления дорожным движением, экологических параметров. В настоящее время мониторинг понимается как измерение интенсивности транспортного потока и средней скорости. Таким образом, отсутствует возможность получения взаимосвязанных оценок параметров транспортной инфраструктуры, что накладывает ограничения на применимость атрибутивной модели и реализующей ее системы.

В предложенной модели не предполагается наличие таких сложных технических средств мониторинга. Предполагается, что дорожная ситуация оценивается водителем и пересчёт маршрута осуществляется «на лету».

#### Алгоритм построения оптимального маршрута транспортной сети.

В задаче моделирования загрузки автодорожной сети можно выделить следующие этапы:

- оценка входных и выходных потоков;
- определение структуры, размерности и компонент матриц смежности
- определение оптимальных маршрутов с учетом загруженности транспортной сети

Для получения оптимизационной модели построения маршрута разработана математическая модель распределения потоков, основанная на 1 и 2 принципах Вардропы. На основе данной модели разрабатывается алгоритм оптимизации транспортных потоков[2].

Пусть необходимо добраться из одной точки транспортной сети в другую с минимальными временными затратами. В качестве модели был выбран ориентированный граф, каждой дуге которого сопоставлены две оценки: длина пути и время задержки (простоя в пробке). Матрица смежности подобного графа имеет вид (1).

Представленный граф может быть описан тензором  $G_{ij}^{lt}$ , где  $l$  - длина пути от  $i$  до  $j$  точки маршрута

$t$  - время задержки на интервале маршрута от  $i$  до  $j$ .

Разработанный алгоритм основан на итеративном нахождении кратчайшего пути с последующим исключением дуги с наибольшим значением временной задержки, найденной на оптимальном пути, и перестроением исходного графа.

Обобщённо разработанный алгоритм можно описать следующим образом:

1. Нахождение кратчайшего пути из вершины  $s$  до вершины  $t$ .

2. Проверка условия допустимости найденного в п.1 пути по параметру временной задержки. По значению “True” – конец алгоритма, иначе переход к п.3.

3. Удаление из найденного в п.1 пути дуги с максимальным значением задержки и переход к шагу 1.

В качестве алгоритма нахождения кратчайшего пути был выбран алгоритм Флойда [7]. Преимущество данного алгоритма заключается, прежде всего, в универсальности, так как рассчитываются кратчайшие пути между всеми возможными парами вершин матрицы смежности. Суть алгоритма заключается в последовательном улучшении пути между парой вершин  $i$  и  $j$  путём замены существующего пути новым, проходящим через промежуточную вершину  $k$ . Во время работы этого алгоритма формируется матрица, которая позволяет восстанавливать пути между требуемыми вершинами. При изменении пропускной способности какой-либо дуги нет необходимости рассчитывать веса дуг для всего графа и повторно выполнять алгоритм.

Результат работы алгоритма Флойда – Уоршелла для заданного  $G(X, V)$  графа можно представить в виде тензора  $G_{ij}^{l_{\min}}$ . Пусть управляющим критерием построения оптимального маршрута будет время задержки  $\tau \leq \tau_{\max}$ , тогда алгоритм построения маршрута будет иметь следующую математическую формулировку:

Шаг 1. Выбор оптимального по параметру минимального расстояния ( $l_{\min}$ ) маршрута ( $\mu_{st}$ ) от вершины  $s$  до вершины  $t$  графа  $G(X, V)$  будет соответствовать операции тензорного пересечения следующего вида:

$$\mu_{st} = G_{ij}^{l_{\min}} \cap G_{st}^{l_{\min}}$$

Шаг 2. Просмотр в исходной матрице, представленной тензором  $G_{ij}^{l_{\tau}}$  параметра времени задержки  $\tau$ :  $\tau_{ij} = G_{ij}^{l_{\tau}} \times \mu_{st} \quad \forall (x_i, x_j) \in \mu_{st}$

Шаг 3. Проверка управляющего критерия времени задержки для всех дуг составляющих маршрут  $\mu_{st}$ :

если  $\tau_{ij} \leq \tau_{\max}$  тогда маршрут  $\mu_{st}$  считается оптимальным, иначе из исходной матрицы удаляются все дуги с задержкой, превышающей  $\tau_{\max}$  и осуществляется переход к шагу 1. Заметим, что в данном алгоритме операция удаления дуги осуществляется посредством замены соответствующего элемента исходной матрицы на бесконечность.

Рассмотрим транспортную сеть, представленную на рисунке 2.

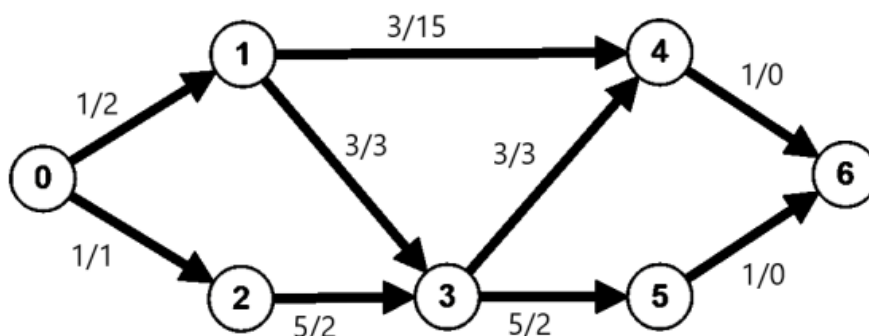


Рисунок 2. Пример графа транспортной сети

Данной сети будет соответствовать матрица смежности следующей структуры:

$$G = \begin{pmatrix} & s & a & b & c & d & t \\ 0 & \infty & \|3 \ 5\| & \|2 \ 2\| & \infty & \infty & \infty \\ 1 & \|3 \ 0\| & \infty & \|1 \ 2\| & \|2 \ 3\| & \infty & \infty \\ 2 & \|2 \ 0\| & \|1 \ 2\| & \infty & \infty & \|3 \ 4\| & \infty \\ 3 & \infty & \|2 \ 0\| & \infty & \infty & \|1 \ 1\| & \|1 \ 2\| \\ 4 & \infty & \infty & \|3 \ 0\| & \|1 \ 0\| & \infty & \|4 \ 5\| \\ 5 & \infty & \infty & \infty & \|1 \ 0\| & \|4 \ 0\| & \infty \end{pmatrix}$$

Пусть требуется найти оптимальный маршрут из узла  $s = 0$  (исток) в узел  $t = 6$  (сток) при  $\tau_{\max} = 3$ .

Согласно алгоритму Флойда кратчайший путь от вершины  $s$  до вершины  $t$   $\mu_{st} = \{0,1,4,6\}$ . Длина данного пути равна  $l_{\min} = 5$ . Согласно второму шагу алгоритма дуга (1,4) имеет временную задержку  $\tau_{ij} > \tau_{\max}$  ( $15 > 3$ ), поэтому данную дугу необходимо убрать из матрицы  $G$ .

Тогда оптимальным будет признан маршрут  $\mu_{st} = \{0,2,3,4,6\}$ , в котором на каждом отрезке выполняется условие  $\tau_{ij} \leq 3$ .

На основе предложенного алгоритма разработан программный модуль. На рисунке 3 представлен результат тестирования программы, реализующей описанный выше подход. Следует отметить, что данная реализация носит экспериментальный характер без претензий на интеграцию в действующую систему навигации транспорта, поэтому интерфейс программы достаточно простой. Однако, программа позволяет не только проложить оптимальный маршрут, который соответствует теоретическим выкладкам, но также рассчитывает среднее время прохождения данного маршрута.

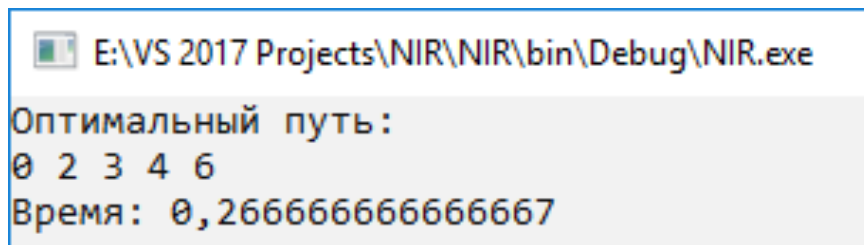


Рисунок 3. Результат работы программы, реализующей представленный алгоритм

### Заключение.

В статье предложена математическая модель матрицы смежности транспортной сети с учетом многомерной структуры учитываемых характеристик и алгоритм нахождения оптимального маршрута. К достоинствам предложенного подхода можно отнести, то, что в условиях постоянного изменения динамических характеристик сети тензорная модель обеспечивает гибкость и возможности адаптации и, следовательно, изменения значений управляющих параметров не будут нести высокой вычислительной нагрузки на алгоритм.

В дальнейшем планируется тщательно исследовать вопрос увеличения скорости путем альтернативных вариантов маршрута и использования методов параллельной реализации предложенного алгоритма.

### Список литературы

[1] Michael Hausenblas, "Applying the Big Data Lambda Architecture", November 12, 2013. Retrieved February 10, 2018, from <http://www.drdoobs.com/database/applying-the-big-data-lambda-architectur/240162604>.

[2] Томас Х. Кормен, Чарльз И. Лейзерсон, Рональд Л. Ривест, Клиффорд Штайн. Алгоритмы: построение и анализ. — 2-е изд. — М.: Вильямс», 2006. — С. 1296.

[3] В.А. Пышный. Моделирование загрузки транспортной сети // Известия ТулГУ. Технические науки. 2012. Вып. 2. С. 457-473.

[4] Гасников А.В., Кленов С.Л., Нурминский Е.А. и др. Введение в математическое моделирование транспортных потоков: учеб. пособие. М.: МФТИ, 2010. — 362 с.

[5] В. И. Швецов. Математическое моделирование транспортных потоков, Авто- мат. и телемех., 2003, выпуск 11, 3-46

[6] Aleksandr Bakulev, Marina Bakuleva, Sergei Skvortsov, Maksim Kozlov, Tatiana Pyurova, Vladimir Hrukin. Modern approaches to the development parallel programs for modern multicore processors.. Proceedings of 6th Mediterranean Conference on Embedded Computing (MECO), Bar, Montenegro, 2017, pp.38-4

[7] Bakulev A.V., Bakuleva M.A., Avilkina S.B. Mathematical methods and algorithms of mobile parallel computing on the base of multi-core processors // European researcher. 2012. V. 33. № 11-1. P. 1826-1834.

## **THE TRANSPORT NETWORK PATH OPTIMIZATION ALGORITHM WITH DYNAMIC PARAMETERS**

***V.P. Koryachko***

*Head of the Department of Computer Aided Design of Computing Facilities Ryazan State Radio Engineering University named after V.F. Utkin (RSREU), Honored Worker of Science and Technology of the Russian Federation, Doctor of Sciences in Engineering Science, Professor*

***A.V. Bakulev***

*Associate Professor of the Department of Computer Aided Design of Computing Facilities Ryazan State Radio Engineering University named after V.F. Utkin (RSREU), PhD of Technical Sciences, Associate Professor*

***M. A. Bakuleva***

*Associate Professor of the Department of Computer Aided Design of Computing Facilities Ryazan State Radio Engineering University named after V.F. Utkin (RSREU), PhD of Technical Sciences, Associate Professor*

*Department of Computer Aided Design of Computing Facilities Ryazan State Radio Engineering University named after V.F. Utkin (RSREU)*

*E-mail: koryachko.v.p@rsreu.ru, alex.bakulev@gmail.com, marina.bakuleva@gmail.com*

**Abstract.** The paper devote to the algorithm of optimization, with the agile functionality, that allow solve the optimization problem considerate with dynamic factors. The control parameters (factors) have different semantics and metrics, for example, travel time and transport network bandwidth. A tensor model is used as a mathematical model that allows to adapted under the external factors. Floyd–Warshell algorithm is proposed as a basic optimization algorithm. This algorithm has been modified in accordance with the research purpose.

**Keywords:** transport network, throughput, tensor model, matrix models of transport network, graphs, algorithms on graphs, optimal path, Floyd algorithm.

УДК 004.021:004.75

## ФИНАНСОВЫЙ НАВИГАТОР: КОНЦЕПТ ПЛАТФОРМЫ ДЛЯ УПРАВЛЕНИЯ ЛИЧНЫМИ ФИНАНСАМИ НА ОСНОВЕ АНАЛИЗА БОЛЬШИХ ДАННЫХ



**Д.Г. Готченя**  
Студент 3 курса,  
кафедры Информатики,  
БГУИР  
lolkanalolke@mail.ru



**О.Э. Осадчий**  
Студент 3 курса,  
кафедры Информатики,  
БГУИР  
trapper.md@gmail.com



**В.Д. Владымцев**  
Ассистент кафедры  
Информатики, инженер-  
программист ОИАСУ ЦИИР  
БГУИР  
v.vladymtsev@bsuir.by

### **Д.Г. Готченя**

Студент 3 курса специальности «Информатика и технологии программирования» БГУИР.

### **О.Э. Осадчий**

Студент 3 курса специальности «Информатика и технологии программирования» БГУИР.

### **В.Д. Владымцев**

Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР.

**Аннотация.** Данная научная работа посвящена разработке платформы для управления личными финансами на основе анализа больших данных. В работе рассматривается методика анализа данных из различных источников, таких как банковские счета, кредитные карты и инвестиционные портфели, а также рассматриваются методы оптимизации личных финансовых решений. Платформа предоставляет пользователю удобный интерфейс для взаимодействия с его личными финансами и помогает принимать обоснованные финансовые решения на основе анализа больших данных. Результаты работы могут быть полезными для людей, желающих эффективно управлять своими финансами и улучшить свое финансовое положение.

**Ключевые слова:** Big data, LLM, большие данные, машинное обучение, финансы, финансовая грамотность, нейронные сети.

### **Введение.**

В настоящее время управление личными финансами является одним из важных аспектов жизни большинства людей. С появлением большого количества источников финансовых данных, таких как банковские счета, кредитные карты, инвестиционные портфели и другие, управление личными финансами стало сложной задачей.

Для решения этой задачи могут быть использованы методы анализа данных и моделирования, а также различные программные платформы и приложения, которые позволяют удобно и эффективно управлять личными финансами.

В данной научной работе рассматривается разработка платформы для управления личными финансами, которая основывается на анализе больших объемов данных и прогнозировании финансовых решений.

Цель работы заключается в создании концепта приложения, которое поможет пользователю легко и удобно управлять своими финансами, а также принимать обоснованные решения в области инвестирования и расходования средств. И в дополнение на основе обрабатываемых данных давать пользователю советы и уроки по финансовой грамотности.

### **Актуальность.**

Управление личными финансами является одним из важнейших аспектов повседневной жизни людей, особенно в условиях постоянного роста объема и разнообразия финансовых данных, доступных для анализа и управления.

Однако, существующие решения не всегда удовлетворяют потребности пользователей в комфортном и интуитивно понятном интерфейсе, а также в точности и надежности прогнозирования финансовых решений. В связи с этим, разработка новых программных платформ и приложений, использующих методы анализа больших данных и моделирования, является актуальной задачей.

Цель данной научной работы заключается в разработке концепта приложения для управления личными финансами, основанного на анализе больших объемов финансовых данных и прогнозировании финансовых решений.

### **Постановка проблемы**

Перед тем как озвучивать свое видение платформы для управления финансами стоит рассмотреть уже существующие на рынке аналоги. Необходимо рассмотреть их сильные и слабые стороны.

На рынке существуют множество приложений, которые помогают пользователям управлять своими финансами, но у них есть и некоторые проблемы.

Первая проблема заключается в неэффективности управления большим объемом данных: с появлением большого количества источников финансовых данных, таких как банковские счета, кредитные карты, инвестиционные портфели и другие, управление личными финансами стало сложной задачей. Некоторые приложения не обеспечивают эффективного управления большим объемом данных, что приводит к трудностям в отслеживании расходов и управлении бюджетом.

К примеру, приложение Mint. Хотя Mint популярен как инструмент для управления финансами, многие пользователи жалуются на то, что приложение не может интегрироваться со всеми их финансовыми учетными записями. Это ограничивает способность Mint эффективно управлять большим объемом данных, что может приводить к неточности при анализе финансовой ситуации.

Вторая проблема заключается в ограниченности функционала: многие приложения для управления финансами имеют ограниченный функционал, который не удовлетворяет потребности пользователей. Они могут не иметь функционала для работы с определенными типами финансовых активов, или не предоставлять достаточно информации для принятия обоснованных финансовых решений. Такого рода недостатки есть у приложения Acorns — это приложение, которое автоматически инвестирует сдачу от покупок пользователя в портфель ETF.

Однако, приложение не предоставляет полной информации о том, как работает инвестиционный портфель, и не позволяет пользователям выбирать конкретные инвестиции. Это ограничение может приводить к тому, что пользователи не имеют достаточной информации для принятия обоснованных инвестиционных решений.

Третья проблема заключается в отсутствии безопасности. Многие приложения собирают и хранят финансовые данные пользователей, но не всегда обеспечивают достаточную защиту данных. Это может привести к утечкам информации или краже личных данных.

Это не все существующие проблемы, однако они являются самыми значимыми, которые необходимо решать в первую очередь

В свою очередь анализируя уже существующие решения, а также цели, которые должна решать программа, возникает список модулей, которые необходимо реализовать:

1. Автоматический анализ трат: Многие существующие приложения уже позволяют пользователям ввести данные о своих финансовых транзакциях, но только некоторые из них могут проанализировать эти данные и предоставить пользователям полезную информацию. Например, приложение Mint может анализировать траты и предоставлять пользователю отчеты, в которых отображаются его расходы по категориям, таким как продукты питания, развлечения и т.д.



2. Управление долгами: Одной из проблем, с которыми сталкиваются многие люди, является управление долгами и кредитами. В приложениях такого рода можно добавить функциональность для отслеживания долгов и кредитов, а также создания планов по их погашению. Например, приложение Debt Manager помогает пользователям отслеживать свои долги, расходы и доходы.

3. Интеграция с банковскими счетами и инвестиционными портфелями: Некоторые приложения уже позволяют пользователям связывать свои банковские счета и инвестиционные портфели с приложением. Таким образом, пользователь может видеть всю свою финансовую информацию в одном месте и производить анализ данных на основе этой информации. Например, приложение Personal Capital позволяет связывать счета и портфели с различных финансовых учреждений, чтобы дать полное представление о финансовом положении пользователя.

4. Рекомендации по инвестированию: Важной функцией платформы для управления личными финансами может стать предоставление пользователю рекомендаций по инвестированию. Некоторые приложения уже предоставляют пользователю рекомендации по инвестированию на основе его финансовых данных и целей. Например, приложение Acorns предлагает пользователю автоматически инвестировать свои деньги в соответствии с целями и рисками, которые пользователь задал.

В анализе доходов можно использовать множество методов анализа больших данных. Но наиболее интересными являются связка технологий *решающих деревьев* и *случайного леса*.

*Решающие деревья и случайный лес* — это методы машинного обучения, которые могут использоваться для анализа данных доходов в финансовых приложениях.

Решающее дерево [1] — это структура данных, которая представляет собой дерево, где каждый узел представляет собой тест на значение одного из признаков данных. Каждый лист дерева соответствует прогнозу по целевой переменной. Решающие деревья могут быть использованы для прогнозирования доходов, например, на основе признаков, таких как возраст, образование и опыт работы. Однако, решающие деревья могут иметь проблемы с переобучением, когда они становятся слишком сложными и точными на обучающих данных, но не могут обобщить свои результаты на новые данные.

Чтобы решить эту проблему, можно использовать случайный лес [2] — это метод ансамбля, который комбинирует несколько решающих деревьев. Каждое дерево обучается на случайной подвыборке данных и случайном наборе признаков. Затем прогнозы каждого дерева комбинируются для получения более точного прогноза.

Случайный лес может использоваться для анализа доходов и расходов, например, для прогнозирования дохода на основе таких признаков, как возраст, образование, опыт работы, регион проживания и другие социально-экономические факторы.

Конкретные примеры использования решающих деревьев и случайного леса в личном финансовом приложении могут включать в себя:

1. Поиск наиболее выгодных предложений: использование решающих деревьев и случайного леса для анализа данных о предложениях от различных банков, страховых компаний и других финансовых учреждений и предоставление пользователю рекомендаций по выбору наиболее выгодных условий. Приложение может анализировать кредитную историю и потребности, и предлагать наиболее подходящие кредитные карты, которые подходят именно этому пользователю. Это может помочь получить лучшую кредитную карту, которая предлагает максимальное количество вознаграждений, но не слишком высокий процент.

2. Автоматическое бюджетирование: использование решающих деревьев для автоматического бюджетирования расходов пользователя. На основе предыдущих расходов и доходов можно определить категории расходов, на которые стоит уменьшить расходы и наоборот, и предоставить рекомендации для максимально эффективного использования доходов.

3. Автоматическое инвестирование. Приложение может использовать машинное обучение, чтобы анализировать портфель инвестиций и давать рекомендации по оптимальным инвестициям, учитывая инвестиционные цели и уровень риска пользователя.

В самом начале, перед тем как начать анализировать траты, совершаемые пользователем, стоит их разбить на группы. С этим хорошо справляется метод классификации.

Классификация является одним из основных методов машинного обучения и используется для разделения объектов на заранее заданные классы на основе их признаков. В задачах анализа трат, классификация может быть использована для определения типа траты на основе ее описания или других признаков.

Существует множество алгоритмов классификации, включая наивный Байесовский классификатор [3], решающие деревья, метод k-ближайших соседей, логистическую регрессию и многие другие. Выбор конкретного алгоритма зависит от особенностей задачи и характеристик обучающего набора данных.

Одним из практических применений классификации в анализе трат, может быть, автоматическое распределение трат на разные категории, например, «питание», «транспорт», «одежда», «развлечения» и т.д. При этом, можно использовать различные признаки для классификации, такие как текстовое описание траты, сумма платежа, дата и место проведения траты и т.д. Так же возможна реализация модуля парсинга необходимой информации по фотографии чека.

После получения набора категорий трат, их можно преобразовать в связанные между собой ячейки, тем самым образовав граф, и уже для него применять анализ графов.

Анализ графов [4] — это метод, который используется для анализа связей между объектами в больших наборах данных. В контексте управления личными финансами, графы могут быть использованы для моделирования финансовых потоков и выявления зависимостей между тратами и доходами.

С помощью анализа графов можно выявлять паттерны потребления, такие как, например, наиболее распространенные категории трат, сезонность расходов, повторяющиеся покупки и т.д.

Данные выводы могут быть использованы для дальнейшей оптимизации бюджета и улучшения финансового здоровья. Кроме того, с помощью анализа графов можно выявлять взаимосвязи между различными категориями трат.

Например, можно выявить, что люди, которые покупают много одежды, склонны также тратить много денег на рестораны. Эта информация может быть использована для того, чтобы давать рекомендации по снижению трат в определенных категориях.

#### **Заключение.**

Итогом данной научной работы разработан концептуальная архитектура финансового помощника. Были подробно рассмотрены существующие аналоги и сделаны выводы на их основе.

Так как развитие в сфере машинного обучения не стоит на месте, то впоследствии из данного концепта может получиться программный продукт, который может стать незаменимым приложением для повседневной жизни обычного человека.

В целом, использование методов машинного обучения в финансовых приложениях может помочь пользователям в оптимизации своих финансовых решений и принятии более обоснованных решений в планировании своих финансов на будущее.

#### **Список литературы**

[1] Решающие деревья – [Электронный ресурс]. Режим доступа: <https://academy.yandex.ru/handbook/ml/article/reshayushchiye-derevya> (Дата обращения: 01.04.2023).

[2] Случайный лес (Random Forest) – [Электронный ресурс]. Режим доступа: <https://alexanderdyakonov.wordpress.com/2016/11/14/случайный-лес-random-forest/> (Дата обращения: 02.04.2023).

[3] Наивный байесовский классификатор – [Электронный ресурс]. Режим доступа: <https://analytics.github.io/data-mining/072-NBC.html> (Дата обращения: 02.04.2023).

[4] М. Свами Графы, сети и алгоритмы / М. Свами. –: ЕЁ Медиа, 2009. – 450 с.

## **FINANCIAL NAVIGATOR: THE CONCEPT OF A PLATFORM FOR PERSONAL FINANCE MANAGEMENT BASED ON BIG DATA ANALYSIS**

***D.G. Gotchenya***

*3rd year student, Department of Informatics, BSUIR*

***A.E. Asadchy***

*3rd year student, Department of Informatics, BSUIR*

***V.D. Vladymtsev***

*Assistant of the Department of Informatics, Enegenier – Programmer Department of Integrated Automated Control Systems, BSUIR*

*Department of Information and Computer Systems Design  
Faculty of Computer Networks and Systems*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: lolkanalolke@mail.ru, trapper.md@gmail.com, v.vladymtsev@bsuir.by*

**Abstract.** This scientific work is devoted to the development of a platform for personal finance management based on big data analysis. The paper discusses the methodology for analyzing data from various sources, such as bank accounts, credit cards and investment portfolios, as well as methods for optimizing personal financial decisions. The platform provides the user with a user-friendly interface for interacting with their personal finances and helps them make informed financial decisions based on big data analysis. The results of the work can be useful for people who want to effectively manage their finances and improve their financial situation.

**Keywords:** Big data, LLM, big data, machine learning, finance, financial literacy, neural networks.

УДК 621.3.049.77–048.24:537.2

## ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ИНТЕРНЕТ-МАРКЕТИНГЕ



**Д.В. Тавлуй**

Магистрант кафедры проектирования информационно-компьютерных систем БГУИР  
dmitriytavlyuy@gmail.com



**В.Ф. Алексеев**

Доцент кафедры проектирования информационно-компьютерных систем БГУИР, кандидат технических наук, доцент  
alexvikt.minsk@gmail.com



**Г.А. Пискун**

Доцент кафедры проектирования информационно-компьютерных систем БГУИР, кандидат технических наук, доцент  
alexvikt.minsk@gmail.com

### **Д.В. Тавлуй**

Окончил БГУИР (2021 г.), в настоящее время является магистрантом этого университета. Проводит научные исследования по алгоритмам и методам интернет-маркетинга в конкурентной среде.

### **В.Ф. Алексеев**

Окончил Минский радиотехнический институт. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

### **Г.А. Пискун**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с моделированием и оптимальным проектированием информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

**Аннотация.** В статье рассматриваются области и аспекты применения искусственного интеллекта в интернет-маркетинге, рассмотрены плюсы и минусы, тенденции развития машинного обучения и его влияния на маркетинговый рынок.

**Ключевые слова:** искусственный интеллект, машинное обучение, анализ, оптимизация.

### **Введение.**

Искусственный интеллект (ИИ) может быть полезным инструментом в интернет-маркетинге. Он может помочь с анализом больших объемов данных, повышением точности прогнозов и улучшением пользовательского опыта.

Одной из основных тенденций применения искусственного интеллекта в интернет-маркетинге является персонализация контента. С помощью ИИ можно анализировать данные о поведении пользователей и создавать персонализированный контент, который будет наиболее интересен конкретному пользователю.

Другой важной тенденцией является автоматизация маркетинговых процессов. С помощью ИИ можно автоматически анализировать данные о поведении пользователей, создавать и оптимизировать рекламные кампании и управлять взаимодействием с клиентами.

В последнее время стали активно развиваться инструменты анализа голоса и обработки естественного языка (NLP). Они позволяют создавать голосовых помощников и автоматически анализировать текстовые данные, что может быть полезно для мониторинга отзывов клиентов и управления репутацией бренда в социальных сетях.

Кроме того, ИИ также используется для улучшения поисковой оптимизации (SEO) и анализа данных, что может помочь компаниям лучше понимать своих клиентов и оптимизировать свои маркетинговые кампании [1-7].

#### **Персонализация контента.**

Персонализация контента с помощью искусственного интеллекта является важным трендом в интернет-маркетинге. С помощью ИИ можно анализировать данные о поведении пользователей на сайте и создавать персонализированный контент, который будет наиболее интересен каждому конкретному пользователю. Например, ИИ может анализировать данные о том, какой контент был просмотрен пользователем, на какие категории товаров он смотрел и что искал на сайте. Используя эти данные, ИИ может предложить пользователю продукты или услуги, которые наиболее подходят под его интересы.

Кроме того, ИИ может также анализировать данные о том, как пользователь взаимодействует с сайтом, например, какие кнопки он нажимает и как долго он остается на странице. Эти данные могут использоваться для оптимизации пользовательского опыта и предложения более удобного интерфейса.

В целом, персонализация контента с помощью ИИ позволяет компаниям улучшить взаимодействие с клиентами и увеличить конверсию, так как каждый пользователь получает предложения и информацию, которые наиболее соответствуют его потребностям и интересам [1].

Для персонализации контента в интернет-маркетинге можно использовать:

1 *Рекомендательные системы.* ИИ может использоваться для создания рекомендательных систем, которые предлагают пользователю наиболее подходящие товары, услуги или контент на основе анализа истории просмотров и поведения на сайте.

2 *Анализ текста.* ИИ может использоваться для анализа текстовых данных, таких как отзывы клиентов, комментарии и социальные медиа-посты, для определения предпочтений и интересов пользователей и создания персонализированного контента.

3 *Определение момента покупки.* ИИ может использоваться для анализа поведения покупателей на сайте и определения момента, когда они готовы совершить покупку. На основе этой информации можно создавать персонализированный контент, который мотивирует пользователя совершить покупку.

4 *Персонализированные E-mail-рассылки.* ИИ может использоваться для создания персонализированных email-рассылок, которые содержат информацию, которая может быть наиболее интересна конкретному пользователю.

Кроме того, использование ИИ для персонализации контента помогает улучшить взаимодействие с клиентами, увеличить лояльность и увеличить конверсию [2].

#### **Анализ поведения пользователей.**

Анализ поведения пользователей может помочь компаниям лучше понимать потребности и интересы своих клиентов и создавать персонализированный контент и предложения, которые увеличивают конверсию и улучшают взаимодействие с клиентами.

Для анализа поведения пользователей ИИ может использовать данные о пользователях. Искусственный интеллект может быть использован для анализа больших объемов данных, связанных с поведением пользователей на сайте или в приложении.

Это может включать анализ данных о просмотрах страниц, посещениях сайта, покупках и других взаимодействиях пользователей с контентом и продуктами. ИИ может быть использован для сегментации пользователей на основе их поведения на сайте или в приложении. Это может помочь компаниям создавать персонализированный контент и предложения для каждого сегмента пользователей, учитывая их интересы и потребности. Искусственный интеллект может быть использован для прогнозирования поведения пользователей на основе анализа их истории взаимодействия с контентом и продуктами. Это может помочь компаниям предсказывать, какие товары или услуги могут быть интересны каждому конкретному пользователю и предлагать им

соответствующий контент. ИИ может быть использован для создания персонализированного контента для каждого пользователя на основе его истории просмотров и поведения на сайте.

Это может включать персонализированные рекомендации товаров, контента или услуг, которые наиболее соответствуют интересам пользователя. ИИ может быть использован для улучшения пользовательского опыта на сайте или в приложении. Например, ИИ может быть использован для определения наиболее эффективных мест для размещения контента на странице или для предложения наиболее удобных путей навигации по сайту [3].

С помощью ИИ можно анализировать данные о поведении пользователей на сайте и определить, как они взаимодействуют с контентом и продуктами.

Примеры использования ИИ для анализа поведения пользователей в интернет-маркетинге:

1 *Анализ поведения на сайте.* ИИ может использоваться для анализа того, как пользователи взаимодействуют с сайтом, например, какие страницы они посещают, как долго они находятся на странице, какие товары они просматривают и какие кнопки нажимают. Эта информация может быть использована для оптимизации пользовательского опыта и улучшения конверсии.

2 *Анализ данных социальных медиа.* ИИ может использоваться для анализа данных социальных медиа, например, для определения, какие сообщения получают большее внимание пользователей, какие темы наиболее обсуждаются и как пользователи реагируют на определенные контенты. Это может помочь компаниям создавать более эффективный контент и улучшать взаимодействие с клиентами.

3 *Анализ отзывов клиентов.* ИИ может использоваться для анализа отзывов клиентов, например, на сайтах с отзывами или на социальных медиа. Это поможет компаниям определить, что нравится и что не нравится клиентам, и как можно улучшить свои продукты и услуги.

4 *Прогнозирование поведения пользователей.* ИИ может использоваться для прогнозирования поведения пользователей на основе анализа их истории просмотров и поведения на сайте. Это помогает компаниям предсказывать, какие товары или услуги могут быть интересны каждому конкретному пользователю и предлагать им соответствующий контент.

Анализ поведения пользователей с помощью ИИ позволяет компаниям более точно понимать потребности и интересы своих клиентов и создавать персонализированный контент и предложения, которые могут улучшить взаимодействие с клиентами и увеличить конверсию.

### **Оптимизация SEO.**

Оптимизация SEO (*англ.* Search Engine Optimization – это комплекс мер по улучшению сайта для его ранжирования в поисковых системах) с помощью искусственного интеллекта может быть достигнута через использование различных инструментов и технологий, таких как машинное обучение и анализ данных.

Одним из основных применений ИИ в SEO является анализ ключевых слов и фраз, используемых в контенте. С помощью машинного обучения ИИ может анализировать огромные объемы данных и определять, какие ключевые слова наиболее популярны среди пользователей в поисковых запросах. Это позволяет оптимизировать контент на сайте с учетом этих ключевых слов и фраз, чтобы увеличить видимость сайта в поисковых результатах.

Другой способ использования ИИ в SEO – это анализ данных о поведении пользователей на сайте. ИИ может анализировать данные о времени пребывания на сайте, количество просмотров страниц, клики и другие показатели, чтобы определить, какие страницы сайта нуждаются в оптимизации и какие факторы влияют на поведение пользователей.

Кроме того, ИИ может помочь в определении наиболее эффективных стратегий по продвижению сайта, используя данные о конкурентах и их техниках SEO. Например, ИИ может использоваться для анализа социальных медиа и определения наиболее эффективных тем и форматов контента [4].

Наконец, ИИ может помочь в оптимизации технических аспектов сайта, таких как скорость загрузки, безопасность и мобильная оптимизация. Использование ИИ может помочь в обнаружении и исправлении технических проблем на сайте, которые могут негативно влиять на его ранжирование в поисковых результатах.

В целом, использование ИИ в SEO может помочь оптимизировать сайт и повысить его видимость в поисковых результатах, что может привести к увеличению трафика и повышению конверсии.

#### **Анализ рынка.**

Анализ рынка с помощью искусственного интеллекта может помочь компаниям и предпринимателям принимать более эффективные решения и получать конкурентные преимущества. Некоторые из способов использования ИИ для анализа рынка включают:

1 *Анализ данных о клиентах.* Использование ИИ для анализа данных о клиентах может помочь компаниям понять их поведение и потребности. Это может помочь в прогнозировании спроса на определенные товары и услуги, определении наиболее востребованных функций и возможности улучшения пользовательского опыта.

2 *Анализ конкурентов.* ИИ может помочь компаниям анализировать данных о конкурентах, чтобы понять их стратегии, продукты, рыночную долю и другие аспекты. Это может помочь компаниям определить свою позицию на рынке и разработать более эффективную стратегию.

3 *Анализ социальных медиа.* Использование ИИ для анализа социальных медиа может помочь компаниям понять, что говорят люди о их бренде, продуктах и услугах. Это может помочь компаниям определить, какие изменения могут быть необходимы, чтобы улучшить свою репутацию и удовлетворить потребности клиентов.

4 *Анализ экономических и финансовых данных.* ИИ может помочь компаниям анализировать экономические и финансовые данные, такие как цены на акции, курсы валют, процентные ставки и другие данные, чтобы предсказать тенденции рынка и разработать более эффективную стратегию инвестирования.

5 *Анализ климата.* Использование ИИ для анализа климата может помочь компаниям понимать, как изменения климата могут влиять на производство, снабжение и спрос на продукты и услуги. Это может помочь компаниям разработать более эффективные стратегии и решения для адаптации к изменениям климата.

В целом, использование ИИ для анализа рынка может помочь компаниям получить более глубокое понимание рынка и принимать более обоснованные решения [5].

#### **Улучшение пользовательского опыта.**

Улучшение пользовательского опыта (*UX*) является одной из ключевых задач веб-разработки и веб-дизайна.

Искусственный интеллект может помочь улучшить пользовательский опыт веб-сайта или приложения. Использование ИИ для персонализации пользовательского опыта может помочь веб-сайту или приложению адаптироваться к потребностям каждого пользователя. Это может включать персонализированные рекомендации, подборку контента, подборку продуктов и услуг, а также более персонализированный дизайн интерфейса.

ИИ может использоваться для автоматизации процессов на сайте или в приложении может помочь упростить использование и улучшить опыт пользователя.

Например, автоматизация процесса оформления заказа может сделать процесс более быстрым и удобным для пользователей.

Использование ИИ для анализа поведения пользователя может помочь понять, как пользователи взаимодействуют с сайтом или приложением, какие элементы интерфейса они предпочитают, какие функции наиболее востребованы и т.д. Эта информация может быть использована для оптимизации пользовательского опыта и создания более удобного интерфейса.

ИИ применяется для улучшения поиска на сайте или в приложении может помочь пользователям быстрее и легче находить нужный контент.

Например, ИИ может использоваться для автоматического распознавания запросов пользователей и предложения соответствующих результатов поиска. Не стоит забывать об автоматическом переводе.

Использование ИИ для автоматического перевода контента на разные языки может помочь пользователям, которые говорят на разных языках, лучше понимать содержание сайта или приложения.

В целом, использование ИИ для улучшения пользовательского опыта может помочь компаниям создать более удобный и привлекательный интерфейс, улучшить удовлетворенность пользователей и повысить конверсию.

### **Прогнозирование продаж.**

Прогнозирование продаж с помощью искусственного интеллекта является очень полезным инструментом для бизнеса, который позволяет предсказывать будущие продажи на основе анализа исторических данных и других факторов.

Существует несколько подходов к прогнозированию продаж с помощью ИИ. Один из них – это использование методов машинного обучения, таких как регрессия, деревья решений и нейронные сети, для анализа исторических данных и создания моделей, которые могут предсказывать будущие продажи на основе различных входных параметров, таких как временной период, клиентские данные и информация о продукте [6].

Другой подход – это использование анализа данных и бизнес-интеллекта для создания моделей, которые могут прогнозировать продажи на основе внутренних и внешних факторов, таких как экономические условия, конкуренты и сезонность.

Прогнозирование продаж с помощью ИИ может быть полезно для бизнеса в нескольких отношениях, включая оптимизацию запасов, управление производством и планирование маркетинговых кампаний. Более точные прогнозы продаж также могут помочь бизнесам принимать более обоснованные решения по поводу распределения ресурсов и инвестиций.

С помощью ИИ-алгоритмов можно анализировать огромные объемы данных о продажах, включая исторические данные о продажах, данные о клиентах, маркетинговые данные и другие факторы, влияющие на продажи. Это позволяет создавать точные прогнозы продаж, что позволяет бизнесу принимать более осознанные решения, такие как планирование производства, управление запасами, установление цен и рекламных бюджетов.

Одним из примеров использования ИИ для прогнозирования продаж является Amazon, который использует ИИ-алгоритмы для прогнозирования продаж и управления запасами.

Это позволяет Amazon управлять своими запасами более эффективно, сокращать расходы и обеспечивать более быструю доставку товаров своим клиентам.

Однако, стоит отметить, что ИИ не является универсальным решением и не может гарантировать 100% точность прогнозов. Для получения наилучших результатов, ИИ должен использоваться в сочетании с опытом и экспертизой людей [7].

### **Заключение.**

Использование ИИ в интернет-маркетинге может повысить эффективность маркетинговых кампаний и улучшить результаты.

Однако, для достижения наилучших результатов необходимо учитывать как технические, так и этические аспекты использования искусственного интеллекта.

### **Список литературы**

[1] Петров, И. Искусственный интеллект в интернет-маркетинге / И. Петров. – Санкт-Петербург, 2018. – 25-28 с.

[2] Тазова, В.Д. Контент-маркетинг, как инструмент продвижения компании в интернете / В.Д. Тазова, И.А. Стоянов. – Красноярск : Красноярский государственный аграрный университет, 2017. – 55-60 с.3. Жлукта, П.А.



Эффективность внедрения облачной сrm-системы / П.А. Жлукта. – Минск : Белорусский государственный университет экономический факультет, 2019. – 78-90 с.

[3] Смирнова, Е.П. Большие данные в интернет-маркетинге: как использовать данные для увеличения продаж / Е.П. Смирнова. – Уфа, 2020. – 45-46 с.

[4] Алексеев, В.Ф. Разработка онлайн платформы оценки и финансирования инновационных проектов / В.Ф. Алексеев, Д.В. Лихачевский, Г.А. Пискун // BIG DATA Advanced Analytics: collection of materials of the fourth international scientific and practical conference, Minsk, Belarus, May 3 - 4, 2018 / editorial board: M. Batura [etc.]. - Minsk, BSUIR, 2018. - P. 398 – 404.

[5] Беляев, Д.П. Интернет-маркетинг и искусственный интеллект: как применять новейшие технологии / Д.П. Беляев. – Минск, 2021. – 145-146 с.

[6] Алексеев, В.Ф. Анализ системы маркетинга на предприятии и её совершенствование с использованием Internet-технологий / В. Ф. Алексеев [и др.] // Современные информационные компьютерные технологии: Сб. науч. ст. в 2ч. Ч.1 - Гродно: ГрГУ, 2008. – С. 118 – 122.

[7] Алексеев, В.Ф. Задачи и методы обработки социально-экономической информации / В.Ф. Алексеев // Современные средства связи: материалы XVI междунар. науч.-техн. конф., 27–29 сентября 2011 г. — Минск: УО ВГКС, 2011.— С. 102.

## **USE OF ARTIFICIAL INTELLIGENCE IN INTERNET MARKETING**

***D.V. Tavlyuy***

*Master student of the department  
design of information and  
computer systems BSUIR*

***V.F. Alexseev***

*Associate Professor of the  
Department of Design of  
Information and Computer  
Systems of BSUIR, PhD of  
Technical Sciences, Associate  
Professor*

***G.A. Piskun***

*Associate Professor of the  
Department of Design of  
Information and Computer  
Systems of BSUIR, PhD of  
Technical Sciences, Associate  
Professor*

*Department of Information and Computer Systems Design*

*Faculty of Computer Engineering*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: dmitriytavlyuy@gmail.com*

**Abstract.** The article discusses the areas and aspects of the use of artificial intelligence in Internet marketing, considers the pros and cons, plans the development of machine learning and its research in the marketing market.

**Keywords:** artificial intelligence, machine learning, analysis, optimization.

УДК 004.021:004.75

## ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ BIG DATA ДЛЯ АНАЛИЗА РЫНКА НЕДВИЖИМОСТИ



**Д.А. Брезин**

Студент 4 курса, факультета компьютерных систем и сетей, специальность информатика и технологии программирования  
[nove123hero@yandex.by](mailto:nove123hero@yandex.by)



**С.Н. Нестеренков**

Декан факультета компьютерных систем и сетей БГУИР, кандидат технических наук, доцент  
[s.nesterenkov@bsuir.by](mailto:s.nesterenkov@bsuir.by)



**А.Н. Марков**

Заместитель начальника ципир, старший преподаватель, кафедры ПОИТ  
[a.n.markov@bsuir.by](mailto:a.n.markov@bsuir.by)

### **Д.А. Брезин**

Студент 4 курса БГУИР, факультета компьютерных систем и сетей, специальность «Информатика и технологии программирования».

### **С.Н. Нестеренков**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с моделями и методами искусственного интеллекта, информационными системами и технологиями, современными технологиями управления разработкой программного обеспечения.

### **А.Н. Марков**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с вычислительными системами, облачными вычислениями, распределенными вычислительными системами, балансировкой нагрузки вычислительных систем.

**Аннотация.** Данная работа посвящена тому как можно использовать Big Data в сфере недвижимости, а также для анализа рынка и оптимизации алгоритмов поиска и рекомендаций недвижимости в веб-приложениях.

**Ключевые слова:** Big Data, использование Big Data в сфере недвижимости, анализ, встроженные системы.

### **Введение.**

Недвижимость небезосновательно считают одной из сложнейших сфер, в которой требуется тщательное внимание и точные расчеты во всех аспектах.

Для внедрения Big Data в сферу недвижимости существовало несколько предпосылок. Одной из первых была потребность в более точных расчетах и оценках. Была также потребность в составлении прогнозов в данной сфере. Также, очень важным остается снижение рисков. Это один из наиболее значительных аспектов рынка недвижимости, и благодаря технологии больших данных с этим можно довольно эффективно справиться. Последней предпосылкой в этом вопросе является желание интегрировать IT в отрасль.

Специалистам в области Big Data удалось выполнить эти задачи, а также постоянно развивать это направление.

Чем больше будет объем достоверной информации, тем лучше для будущего развития недвижимости. Алгоритмы прогнозирования позволяют выбирать более качественную информацию и повышают точность расчетов. Использование больших данных уже значительно помогает в оценке девелоперской и ипотечной отраслей, более эффективно урегулировании вопросов страхования и оценке покупательной способности с учетом определенных переменных [1].

### **Актуальность.**

Анализ больших данных – это то, с чем сегодня сталкиваются компании, ведущие свою деятельность практически в любой сфере: маркетинг, ритейл, медицина, рынок недвижимости и т. д. Везде, где есть необходимость обрабатывать большие массивы информации, используются данные технологии, и востребованы специалисты.

Результаты анализа применяются для принятия стратегических решений, продвижения продукции и услуг, в социально значимых проектах и программах. В сфере недвижимости Big Data играет важную роль в составлении прогнозов и снижении рисков в данной сфере.

### **Использование Big Data в сфере недвижимости**

Большие данные меняют традиционные способы работы в сфере недвижимости. Новые технологии помогают автоматически собирать большие объёмы информации в кратчайшие сроки.

Собрав данные о взаимодействии пользователей с сайтами агентств недвижимости и застройщиками, результаты соцопросов, статистику о городском населении, план транспортной системы, можно предсказать потребности клиентов в различных сегментах недвижимости в будущем. Также с помощью технологий Big Data можно анализировать данные социальных сетей для сбора информации о предпочтениях и интересах пользователей, которые могут быть связаны с выбором объектов недвижимости.

Например, анализ постов поможет определить наиболее популярные районы для каждой группы пользователей. Можно собирать данные из множества источников о покупке и продаже объектов недвижимости в конкретном районе, о трафике, демографической ситуации, затем анализировать и сопоставлять эти данные, чтобы предоставить консолидированное представление о тенденциях рынка и ценности недвижимости в том или ином районе [2].

Всё это позволяет инвесторам и арендаторам, которые ищут помещение, более полно и точно оценить возможные риски или определить дальнейшую инвестиционную стратегию.

Однако стоит учитывать, что большие объёмы данных могут оказаться непригодными из-за недостатка ресурсов или низкого качества информации, а также анализ может быть затруднителен из-за отсутствия стандартных форматов данных и различных способов хранения данных.

Решение такого рода проблем поможет улучшить качество анализа данных, а также повысит эффективность использования Big Data в недвижимости.

Для внедрения Big Data в приложение для поиска объектов недвижимости нужно сделать следующие шаги:

1) Необходимо собрать данные о взаимодействиях пользователей с сайтами агентств недвижимости и застройщиками, результаты соцопросов, статистику о городском населении, план транспортной системы, а также данные о покупке и продаже объектов недвижимости, и цен в конкретном районе. Данные социальных сетей о предпочтениях и интересах пользователей, связанные с выбором объектов недвижимости.

2) После сбора данных их необходимо обработать и преобразовать в формат, который можно легко использовать в алгоритмах машинного обучения. В обработку включается очистка данных, устранение дубликатов, устранение ошибок. Для этого могут использоваться следующие инструменты Apache Spark, Apache Hadoop MapReduce, Apache Flink.

3) Нужно проанализировать данные используя машинное обучение или статистический анализ. Целью анализа является определение закономерностей и тенденций на рынке недвижимости.

4) На основе результатов анализа данных нужно разработать алгоритмы поиска и рекомендаций.

5) После разработки алгоритмов поиска и рекомендаций, необходимо их добавить в приложение, а после протестировать на эффективность и точность.

### **Заключение.**

В заключение можно отметить, что Big Data может оказывать существенное влияние на маркетинг и помогает в оценке девелоперской, и ипотечной отраслей. Благодаря Big Data мы можем выяснить, где лучше строить дом, какие новые удобства привлекают покупателей, чего нового хотят клиенты, а также поможет уменьшить затраты и поднять размер прибыли.

### **Список литературы**

[1] Идеальное сочетание больших данных и недвижимости [Электронный ресурс]. URL: <https://businessyield.com/ru/real-estate/benefits-of-big-data-for-real-estate/> (Дата обращения: 15.04.2022).

[2] Агентства скоро могут исчезнуть. Как Property Technologies меняют сферу недвижимости [Электронный ресурс]. URL: <https://probusiness.io/markets/3851-agentstva-skoro-mogut-ischeznut-kak-property-technologies-menyayut-sferu-nedvizhimosti.html> (Дата обращения: 15.04.2022).

[3] Big Data в недвижимости [Электронный ресурс]. URL: <https://gmk.ru/blog/big-data> (Дата обращения: 15.04.2022).

[4] Предсказание цен на недвижимость с помощью Big Data [Электронный ресурс]. URL: [https://taskon.com/blog/predskazaniye\\_tsen\\_big\\_data](https://taskon.com/blog/predskazaniye_tsen_big_data) (Дата обращения: 15.04.2022).

## **USING BIG DATA TECHNOLOGIES TO ANALYZE THE REAL ESTATE MARKET**

***D.A.Brezin***

*4th year student, Faculty of Computer Systems and Networks, specialty Computer Science and Programming Technology*

***S.N.Nesterenkov***

*Dean of the Faculty of Computer Systems and Networks of BSUIR, Candidate of Technical Sciences, Associate Professor*

***A.N. Markov***

*Deputy Head of the CIIR, Senior lecturer, Department of SFIT*

*Department of Computer Science*

*Faculty of computer system and networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: nove123hero@yandex.by*

**Annotation.** This paper is devoted to how Big Data can be used in real estate, as well as for market analysis and optimization of algorithms for searching and recommending real estate in web applications.

**Keywords:** Big Data, the use of Big Data in real estate, analysis, embedded systems.

УДК 004.04.4

## **ПРИМЕНЕНИЕ BIG DATA В УПРАВЛЕНИЕ ПРОТЕЗОМ СЕРДЦА ЧЕЛОВЕКА**



*А.М. Тургунов*

*Профессор кафедры математики и  
естественных наук Каршинского института  
ирригации и агротехнологий НИЦ «ТИИМСХ»  
кандидат технических наук  
adilbekturgunov@gmail.com*

### **А.М. Тургунов**

*Окончил Ташкентский политехнический институт. Область научных интересов связана с разработкой методов и алгоритмов построение бионических протезов человека, применение искусственного интеллекта в системе здравоохранения и организацией учебного и научно-исследовательских работ в техническом институте.*

**Аннотация.** В статье рассматривается применение Big Data в управлении протезом сердца человека. Авторы обсуждают преимущества использования больших данных в данной области, такие как мониторинг сердечной деятельности, анализ данных, прогнозирование осложнений, информирование пациента и защита данных. Однако, необходимы строгие меры по защите данных, чтобы гарантировать конфиденциальность пациента и защитить их от кибератак. В целом, применение Big Data в управлении протезом сердца – это пример использования передовых технологий для улучшения здравоохранения и повышения качества жизни людей.

**Ключевые слова:** Big data, управление протезом сердца, медицинская база данных, машинное обучение, анализ данных, информационные технологии, кардиология, здравоохранение.

### **Введение.**

Сердечно-сосудистые заболевания являются одним из главных вызовов современной медицины и становятся все более распространенными в мире. Управление протезом сердца является одним из ключевых методов лечения сердечной недостаточности, которое может увеличить продолжительность и качество жизни пациентов. Однако, управление протезом сердца требует точной диагностики, мониторинга и оптимального лечения.

В настоящее время, с развитием информационных технологий и возможностей анализа больших объемов данных, применение Big Data в медицине становится все более актуальным. В данной статье мы рассмотрим применение Big Data в управлении протезом сердца человека, что может привести к улучшению результатов и качества жизни пациентов.

Протезы сердца являются жизненно важными устройствами для людей с серьезными заболеваниями сердца. Однако, управление протезом может быть сложным процессом, который требует постоянного мониторинга и анализа данных. В настоящее время, существует возможность использовать технологии Big Data для управления протезом сердца.

Big Data – это совокупность данных, которые собираются из разных источников, и которые могут быть использованы для анализа и принятия решений. В случае управления протезом сердца, данные, полученные от мониторинга сердечной деятельности, могут быть использованы для анализа и принятия решений в режиме реального времени.

Применение Big Data в управлении протезом сердца может значительно повысить эффективность устройства и улучшить качество жизни пациента. Вот некоторые из способов, которыми Big Data может быть применена в управлении протезом сердца.

В данной статье будут обсуждены вопросы, как Big Data анализ данных может помочь медицинским учреждениям и специалистам в сборе и анализе больших объемов данных, таких как данные мониторинга пациента, данные диагностики и обследования, данные о лекарствах и т.д. Кроме того, будет рассмотрена роль машинного обучения и анализа данных в управлении протезом сердца и оптимизации лечения.

Таким образом, данная статья может заинтересовать медицинских специалистов, разработчиков информационных технологий и всех, кто интересуется технологическим прогрессом в здравоохранении.

#### **Актуальность.**

Применение Big Data в управлении протезом сердца человека связана с быстро развивающейся технологией в области медицины и информационных технологий. Сердечно-сосудистые заболевания являются одним из главных причин смертности в мире. Управление протезом сердца является одним из ключевых методов лечения сердечной недостаточности, и применение Big Data может улучшить его эффективность и результаты.

С помощью Big Data анализа данных, медицинские учреждения и специалисты могут собирать и анализировать большие объемы данных, такие как данные мониторинга пациента, данные диагностики и обследования, данные о лекарствах и т.д. Это позволяет увеличить точность диагностики и прогнозирования, оптимизировать лечение и снизить риски для пациентов.

Таким образом, статья о применении Big Data в управлении протезом сердца человека имеет большую актуальность в контексте современного здравоохранения, и может привлечь внимание медицинских специалистов и разработчиков технологий.

Мониторинг сердечной деятельности Big Data может быть использован для мониторинга сердечной деятельности пациента с помощью сенсоров и медицинских устройств. Эти данные могут быть собраны и анализированы в реальном времени, что позволяет врачам быстро выявлять любые неисправности в работе протеза и принимать необходимые меры.

Анализ данных Big Data может быть использован для анализа данных, собранных от протеза сердца. Эти данные могут быть использованы для определения оптимальных параметров работы протеза, таких как скорость и ритм сердечных сокращений. Это может помочь улучшить эффективность протеза и уменьшить риск осложнений.

Прогнозирование осложнений Big Data может быть использован для прогнозирования возможных осложнений, связанных с работой протеза сердца. С помощью алгоритмов машинного обучения, Big Data может анализировать данные о состоянии пациента и работе протеза, и предсказывать возможные проблемы до их возникновения.

Информирование пациента Big Data также может использоваться для информирования пациента о состоянии его сердечной деятельности и работе протеза. Это может помочь пациенту лучше понимать состояние своего здоровья и принимать более обоснованные решения в отношении своего образа жизни. Кроме того, эта информация может быть передана врачу для более точного диагноза и лечения.

Безопасность данных Собранные данные о здоровье пациента являются чрезвычайно конфиденциальными и могут стать объектом кибератак. Big Data может помочь защитить эти данные путем установки протоколов безопасности, таких как шифрование и аутентификация.

В заключение, использование Big Data в управлении протезом сердца может значительно улучшить качество жизни пациентов, а также облегчить работу медицинского персонала. Мониторинг сердечной деятельности, анализ данных, прогнозирование осложнений, информирование пациента и защита данных – все эти функции могут быть эффективно выполнены с помощью Big Data. Однако, необходимы строгие меры по защите данных, чтобы гарантировать конфиденциальность пациента и защитить их от кибератак. В целом, применение Big Data в управлении протезом сердца – это пример использования передовых технологий для улучшения здравоохранения и повышения качества жизни людей.

**Концептуальная модель включает следующие элементы:**

**Сенсоры.** Устройства, которые мониторят работу протеза сердца и собирают данные о сердечной деятельности пациента.

**База данных.** Хранилище, в котором хранятся данные, собранные с помощью сенсоров. База данных может быть расположена на удаленном сервере или на устройстве пациента.

**Аналитика данных.** Алгоритмы, которые обрабатывают и анализируют данные, собранные с помощью сенсоров. Аналитика данных может включать в себя машинное обучение, статистический анализ, моделирование и т.д.

**Пользовательский интерфейс.** Интерфейс, который предоставляет пациенту и медицинскому персоналу доступ к данным о сердечной деятельности. Пользовательский интерфейс может быть в виде приложения для мобильного устройства или веб-интерфейса.

**Система управления.** Система, которая использует данные о сердечной деятельности для управления протезом сердца. Система управления может включать в себя алгоритмы регулирования сердечного ритма, контроль за подачей лекарств и т.д.

**Система безопасности.** Система, которая обеспечивает безопасность данных о сердечной деятельности пациента. Система безопасности может включать в себя механизмы аутентификации, шифрования и контроля доступа.

Концептуальная модель показывает, как различные элементы системы могут взаимодействовать друг с другом для обеспечения управления протезом сердца и мониторинга здоровья пациента. Большие данные играют важную роль в этой системе, обеспечивая сбор, хранение, анализ и использование информации о сердечной деятельности пациента.

Структура базы данных для управления протезом сердца может варьироваться в зависимости от конкретной реализации системы и используемых технологий. Однако, в общих чертах, база данных может включать следующие таблицы:

**Таблица данных о пациенте.** В этой таблице содержится информация о пациенте, такая как ФИО, дата рождения, контактная информация, а также информация о медицинской истории пациента.

**Таблица данных о протезе сердца.** В этой таблице содержится информация о протезе сердца, такая как производитель, модель, серийный номер, дата установки, дата последнего технического обслуживания и т.д.

**Таблица данных о сенсорах.** В этой таблице содержится информация о сенсорах, которые используются для мониторинга сердечной деятельности пациента, такая как тип сенсора, серийный номер, дата установки, дата последней калибровки и т.д.

**Таблица данных о сердечной деятельности.** В этой таблице содержится информация о сердечной деятельности пациента, собранная с помощью сенсоров. Данные могут включать в себя информацию о частоте сердечных сокращений, артериальном давлении, уровне кислорода в крови и т.д.

**Таблица данных об использовании лекарств.** В этой таблице содержится информация об использовании лекарств пациентом, такая как наименование лекарства, дозировка, частота приема, дата начала и окончания приема и т.д.

**Таблица данных о настройках системы управления.** В этой таблице содержится информация о настройках системы управления протезом сердца, такая как алгоритмы регулирования сердечного ритма, настройки скорости и т.д.

**Таблица данных о доступе к базе данных.** В этой таблице содержится информация о пользователях, которым предоставлен доступ к базе данных, а также информация об уровне доступа и правах пользователей.

Структура базы данных должна обеспечивать эффективное хранение, быстрый доступ и защиту данных о пациенте. Важно также учитывать требования к конфиденциальности и защите персональных данных пациента в соответствии с местными законодательными нормами и правилами медицинской этики.

Программный код на Python для создание базы данных "управление протеза сердца человека"

Ниже приведен разработанный программный код для создания базы данных "управление протезом сердца человека":

```
import sqlite3
# Создание базы данных
conn = sqlite3.connect('heart_prosthesis.db')
# Создание таблицы данных о пациенте
conn.execute('''CREATE TABLE patient
              (id INTEGER PRIMARY KEY AUTOINCREMENT,
               full_name TEXT NOT NULL,
               date_of_birth DATE NOT NULL,
               contact_info TEXT,
               medical_history TEXT)''')

# Создание таблицы данных о протезе сердца
conn.execute('''CREATE TABLE heart_prosthesis
              (id INTEGER PRIMARY KEY AUTOINCREMENT,
               manufacturer TEXT NOT NULL,
               model TEXT NOT NULL,
               serial_number TEXT NOT NULL,
               date_installed DATE NOT NULL,
               last_maintenance DATE)''')

# Создание таблицы данных о сенсорах
conn.execute('''CREATE TABLE sensors
              (id INTEGER PRIMARY KEY AUTOINCREMENT,
               type TEXT NOT NULL,
               serial_number TEXT NOT NULL,
               date_installed DATE NOT NULL,
               last_calibration DATE)''')

# Создание таблицы данных о сердечной деятельности
conn.execute('''CREATE TABLE heart_activity
              (id INTEGER PRIMARY KEY AUTOINCREMENT,
               patient_id INTEGER NOT NULL,
               sensor_id INTEGER NOT NULL,
               heart_rate INTEGER NOT NULL,
               blood_pressure INTEGER NOT NULL,
               oxygen_level INTEGER NOT NULL,
               FOREIGN KEY(patient_id) REFERENCES
patient(id),
               FOREIGN KEY(sensor_id) REFERENCES
sensors(id))''')

# Создание таблицы данных об использовании лекарств
conn.execute('''CREATE TABLE medication_usage
              (id INTEGER PRIMARY KEY AUTOINCREMENT,
               patient_id INTEGER NOT NULL,
               medication_name TEXT NOT NULL,
```



```
        dosage TEXT NOT NULL,
        frequency TEXT NOT NULL,
        start_date DATE NOT NULL,
        end_date DATE,
        FOREIGN      KEY(patient_id)      REFERENCES
patient(id))''')

# Создание таблицы данных о настройках системы управления
conn.execute('''CREATE TABLE system_settings
              (id INTEGER PRIMARY KEY AUTOINCREMENT,
               heart_prosthesis_id INTEGER NOT NULL,
               algorithm TEXT NOT NULL,
               speed TEXT NOT NULL,
               FOREIGN KEY(heart_prosthesis_id) REFERENCES
heart_prosthesis(id))''')

# Создание таблицы данных о доступе к базе данных
conn.execute('''CREATE TABLE database_access
              (id INTEGER PRIMARY KEY AUTOINCREMENT,
               user_id INTEGER NOT NULL,
               access_level TEXT NOT NULL,
               permissions TEXT NOT NULL,
               FOREIGN KEY(user_id) REFERENCES users(id))''')

conn.close()
```

Эта программа создает базу данных с названием "heart\_prosthesis.db" и создает таблицы для каждого из пунктов, которые были выше упомянуты. Значения каждого поля могут быть изменены в зависимости от требований проекта.

Для организации запросов в базу данных "управление протеза сердца человека" разработана программа, код которого приведена ниже:  
Получение всех данных о пациентах:

```
import sqlite3
conn = sqlite3.connect('heart_prosthesis.db')
cursor = conn.cursor()
cursor.execute('SELECT * FROM patient')
data = cursor.fetchall()
for row in data:
    print(row)
conn.close()
Получение списка всех протезов сердца определенного
производителя:
python
Copy code
import sqlite3

conn = sqlite3.connect('heart_prosthesis.db')
cursor = conn.cursor()
```

```
manufacturer = 'XYZ Company'
cursor.execute('SELECT * FROM heart_prosthesis WHERE
manufacturer=?', (manufacturer,))
data = cursor.fetchall()

for row in data:
    print(row)

conn.close()
Получение данных о последней калибровке всех сенсоров:
python
Copy code
import sqlite3
conn = sqlite3.connect('heart_prosthesis.db')
cursor = conn.cursor()
cursor.execute('SELECT serial_number, MAX(last_calibration) FROM
sensors GROUP BY serial_number')
data = cursor.fetchall()
for row in data:
    print(row)

conn.close()
```

Обновление данных о дате последнего обслуживания определенного протеза сердца будет осуществлена с помощью следующей программы::

```
import sqlite3

conn = sqlite3.connect('heart_prosthesis.db')
cursor = conn.cursor()

prosthesis_id = 1
new_last_maintenance_date = '2022-01-01'

cursor.execute('UPDATE heart_prosthesis SET last_maintenance=?
WHERE id=?', (new_last_maintenance_date, prosthesis_id))
conn.commit()

conn.close()
```

### **Заключение.**

Применение Big Data в управлении протезом сердца человека может улучшить результаты лечения и качество жизни пациентов. Большие объемы данных, которые собираются в ходе мониторинга и диагностики, могут быть анализированы при помощи машинного обучения и других методов анализа данных, что может привести к более точной диагностике, оптимальному лечению и предотвращению осложнений.

В данной статье мы рассмотрели структуру базы данных для управления протезом сердца, программный код на Python для создания и запросов в базу данных, а также методы анализа данных и машинного обучения для оптимизации лечения и управления протезом сердца.

Применение Big Data в медицине является мощным инструментом для улучшения здравоохранения и лечения пациентов. Однако, необходимо учитывать проблемы конфиденциальности и безопасности данных при их сборе и использовании. Поэтому, необходимо разрабатывать соответствующие меры безопасности и правила использования данных для защиты прав пациентов и сохранения конфиденциальности.

Таким образом, мы надеемся, что данная статья поможет повысить осведомленность о применении Big Data в управлении протезом сердца и стимулировать дальнейшие исследования в этой области, которые могут привести к улучшению здравоохранения и качества жизни пациентов.

#### **Список литературы**

- [1]. Talbot, S. G., & Goldstein, N. E. (2019). The role of big data in the management of heart failure. *Current heart failure reports*, 16(1), 12-20.
- [2] Ziaician, B., & Heidenreich, P. A. (2017). Clinical effectiveness of remote monitoring devices in managing heart failure. *Expert review of medical devices*, 14(11), 907-916.
- [3] Wacker, J. G., Donner-Banzhoff, N., & Krones, T. (2018). Shared decision making and big data in health care. *Journal of medical ethics*, 44(7), 465-470.
- [4] Krittanawong, C., Zhang, H., Wang, Z., Aydar, M., & Kitai, T. (2019). Artificial intelligence in precision cardiovascular medicine. *Journal of the American College of Cardiology*, 69(21), 2657-2664.
- [5] Lee, J. J., Kim, Y. J., Cho, J. H., & Kim, M. S. (2017). Big data analysis for cardiovascular disease prediction using health insurance review and assessment service data. *Health Inform Res*, 23(4), 245-251.
- [6] Johnson, K. W., Torres, Soto J., Glicksberg, B. S., Shameer, K., Miotto, R., Ali, M., ... & Dudley, J. T. (2018). Artificial intelligence in cardiology. *Journal of the American College of Cardiology*, 71(23), 2668-2679.
- [7] Shashikumar, S. P., Shah, A. J., Clifford, G. D., & Nemati, S. (2018). Monitoring and interpreting ICU data using physiological models and machine learning. *IEEE Journal of biomedical and health informatics*, 22(3), 853-860.
- [8] Lupiáñez-Villanueva, F., & Mayer, M. A. (2018). Opportunities and challenges of big data for value-based insurance design. *Frontiers in public health*, 6, 133.
- [9] Demir, Ö. E., & Aydin, K. (2019). Big data analytics in healthcare: A systematic literature review and roadmap for future research. *Journal of medical systems*, 43(8), 233., 138-155.

## **APPLICATION OF BIG DATA IN THE MANAGEMENT OF A HUMAN HEART PROSTHESIS**

***A.M. Turgunov***

*Professor of the Department of Mathematics and Natural Sciences at the Karshi Institute of Irrigation and Agrotechnologies of the NRC "TIAME", Candidate of Technical Sciences.*

*Department of Mathematics and Natural Sciences at the Karshi Institute of Irrigation and Agrotechnologies of the NRC "TIAME", Karshi, Uzbekistan  
E-mail: adilbekturgunov@gmail.com*

**Abstract.** The article examines the application of Big Data in the management of a human heart prosthesis. The authors discuss the advantages of using big data in this field, such as monitoring of heart activity, data analysis, complication forecasting, patient information, and data protection. However, strict measures are needed to protect patient data to guarantee confidentiality and protect them from cyberattacks. Overall, the application of Big Data in heart prosthesis management is an example of the use of advanced technologies to improve healthcare and enhance people's quality of life.

**Keywords:** Big data, heart prosthesis management, medical database, machine learning, data analysis, information technology, cardiology, healthcare.

УДК 004.67:004.75

## СИСТЕМА МОНИТОРИНГА ПОКАЗАТЕЛЕЙ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ДЛЯ ОБРАЗОВАТЕЛЬНЫХ ПЛАТФОРМ НА БАЗЕ MOODLE LMS



**Е.Н. Шнейдеров**

Проректор по учебной работе,  
доцент кафедры проектирования  
информационно-компьютерных  
систем БГУИР  
[shneiderov@bsuir.by](mailto:shneiderov@bsuir.by)



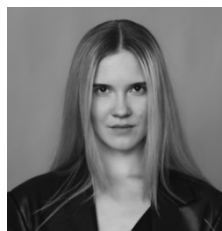
**М.П. Батура**

Доктор технических наук, профессор,  
академик Международной академии  
наук высшей школы  
[bmpbel@bsuir.by](mailto:bmpbel@bsuir.by)



**А. С. Терешкова**

Заместитель начальника центра  
развития дистанционного  
образования БГУИР  
[tereshkova@bsuir.by](mailto:tereshkova@bsuir.by)



**К.С. Крез**

Магистрант БГУИР  
[k.krez@bsuir.by](mailto:k.krez@bsuir.by)

### **Е.Н. Шнейдеров**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой феноменологических моделей и научно-методических основ компьютерного проектирования радиоэлектронных средств, технического обеспечения безопасности и создания электронных систем безопасности.

### **М.П. Батура**

Окончил Минский радиотехнический институт. Область научных интересов связана с системным анализом, управлением и обработки информации в технических и организационных системах

### **А.С. Терешкова**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана со сбором и анализом образовательных данных.

### **К.С. Крез**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов: корреляционный анализ цифровых следов пользователей, нейронные сети, разработка и анализ структуры хранения данных.

**Аннотация.** Разработана система мониторинга показателей образовательного процесса для образовательных платформ на базе Moodle LMS. Описана структурная схема системы, а также возможности ее использования на примере образовательного процесса в системе электронного обучения БГУИР.

**Ключевые слова:** мониторинг образовательного процесса, анализ образовательных данных, Moodle LMS.

Цифровизация является глобальным процессом, охватывающим все сферы общественной жизни, в том числе и сферу высшего образования. Цифровая образовательная среда в учреждениях высшего образования становится полноценным пространством для формирования представления о профессиональном будущем и продуктивных стратегиях поведения [1]. Как правило, цифровая образовательная среда может включать в себя формализованные (предварительно подготовленные) и неформализованные (сформированные случайным образом в этой среде) источники знаний: от электронных материалов учебных дисциплин учреждения образования и индивидуальных преподавательских курсов до обсуждений вопросов и задач на форумах и в личных пользовательских сообщениях. Управление такой цифровой средой и её техническая поддержка является нетривиальной задачей для администраторов информационных систем. Излишние ограничения и установка жёстких правил использования цифровой образовательной среды на практике приводит к значительному замедлению её развития и снижению мотивации пользователей при работе с ней. С другой стороны отсутствие представления о процессах и показателях её функционирования могут повлечь за собой как технические проблемы (например, использование образовательных платформ, не предназначенных для хранения мультимедийных материалов, в целях обмена видео), так и возникновение попыток нецелевого использования образовательных сред.

Целью любой цифровой образовательной среды учреждения высшего образования является быстрый доступ пользователей к информации, необходимой в процессе приобретения обучающимися профессиональных и иных компетенций, предусмотренных образовательной программой. В основе цифровой образовательной среды лежит программное обеспечение для обмена этой информацией (хотя эти понятия и не является тождественными), которое с использованием общедоступных каналов связи позволяет пользователям независимо от места их нахождения работать с учебными материалами. Подобное программное обеспечение называют системами управления обучением или системами электронного (цифрового, дистанционного) обучения (далее – СЭО), включающие в себя ряд технологических возможностей и инструментов для реализации удалённой работы и позволяющие учреждениям образования реализовывать различные схемы образовательного процесса в цифровой среде, в том числе полноценное онлайн-обучение. Статистически наиболее часто используемым учреждениями высшего образования программным обеспечением для реализации образовательного процесса в цифровой образовательной среде является Moodle LMS.

В процессе функционирования СЭО каждый пользователь оставляет в ней так называемый «цифровой след». Цифровой след представляет собой, как правило, информацию о самом пользователе, различные записи в базах данных (о деятельности, о сообщениях, об оценках и др.), отдельные файлы и другие данные. В неявном виде цифровой след пользователя также включает информацию о методике его работы с учебным контентом, а значит, методике его обучения применительно к конкретной дисциплине. Анализ и формализация этой информации может способствовать разработке модели цифрового следа, используя которую можно разработать наиболее оптимальный подход к формированию цифровой образовательной среды [2].

Безусловно, системы электронного обучения, реализованные на базе Moodle LMS, имеют некоторый функционал для мониторинга и анализа организованных с их использованием процессов. Этот функционал основывается на цифровых следах пользователей, хранящихся в базе данных, и реализован на уровне отдельного курса или СЭО в целом. Вместе с тем, в ряде статей [3, 4, 5 и др.] содержится определённый посыл, заключающийся в непригодности и неудобстве универсальных инструментов Moodle LMS для целевого мониторинга деятельности отдельных групп пользователей (например, студентов в группе риска по успеваемости), осуществляемым учреждениями высшего образования. Например, получить достоверные данные о результативности применения СЭО в образовательном процессе на уровне отдельного преподавателя, обучающегося или кафедры невозможно стандартными функциями Moodle

LMS. Дополнительные плагины Moodle LMS сторонних разработчиков также не позволяют комплексно решать задачи мониторинга образовательного процесса.

В учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» с весны 2022 года в рамках выполнения НИР и реализации экспериментального проекта разрабатывается микросервисная программная система мониторинга показателей образовательного процесса в СЭО. Архитектура системы соответствует хранилищам данных с архитектурой шины Ральфа Кимболла [6]. В настоящий момент в ЦОД БГУИР в режиме тестовой эксплуатации развёрнуты 4 виртуальных сервера (рисунок 1), 3 из которых составляют разрабатываемую систему:

- LMS Server (12 CPU / 16 Gb RAM) – непосредственно СЭО на базе Ubuntu 20.04 / Moodle 4.1 с хранилищем на базе MariaDB (не относится к системе мониторинга показателей образовательного процесса);
- ETL Server (4 CPU / 4 Gb RAM) – сервер ETL-инструмента для миграции данных из СЭО в хранилище на базе Ubuntu 22.04 / Apache NiFi 1.20;
- DWH Server (4 CPU / 4 Gb RAM / СХД с расширяемым дисковым пространством) – хранилище данных на базе Ubuntu 22.04 / PostgreSQL 12;
- Monitor Server (4 CPU / 4 Gb RAM) – сервер визуализации данных на базе Ubuntu 20.04 / Apache Superset 2.0.

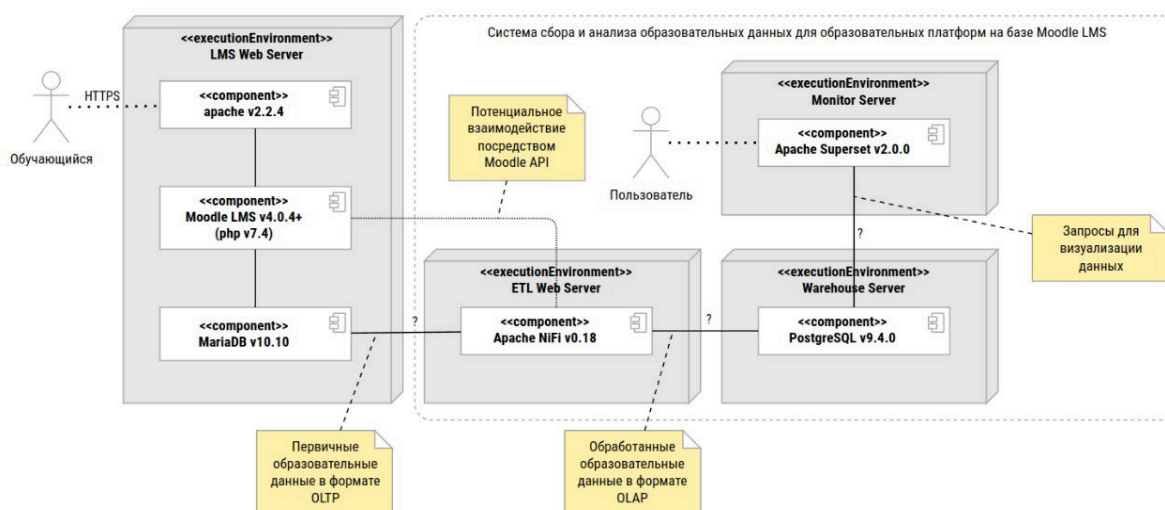


Рисунок 1. Структурная схема системы сбора и анализ данных для образовательной платформы на базе Moodle LMS

За календарный год собрано порядка 2 Тб содержательных и поведенческих цифровых следов пользователей (результат деятельности в СЭО студентов всех форм получения образования и профессорско-преподавательского состава БГУИР), которые в виде предопределённых авторами моделей в реляционном формате хранятся на DWH Server.

Для миграции образовательных данных пользователей с LMS Server (базы данных СЭО) в DWH Server авторы используют бесплатный ETL-инструмент Apache NiFi, имеющий открытый исходный код. Его использование позволяет минимизировать количество скриптов миграции, тем самым упрощая и ускоряя реализацию процесса, сэкономя ресурсы на его поддержку и снизив стоимость разработки. Взаимодействие с базами данных осуществляется посредством SQL-запросов. Apache NiFi имеет встроенные утилиты для:

- взаимодействия с различными источниками данных;

- трансформации данных (валидации, очистки, изменения структуры);
- автоматизации процессов (создания триггеров, планировщиков и др.).

Визуализация процесса миграции цифровых следов пользователей в Apache NiFi представляет собой алгоритм, узлы которого связаны друг с другом потоком файлов. Каждый из узлов принимает или генерирует один входной файл, при необходимости изменяет информацию в нём, после чего передаёт результат дальше по алгоритму. Такое представление позволяет локализовать ошибки на уровне отдельных узлов, выполняющих отдельную функцию. В Apache NiFi узлы алгоритма представляют собой так называемые «процессоры».

Пример такого алгоритма, реализующего миграцию изменений элементов дисциплин в СЭО, приведён на рисунке 2.

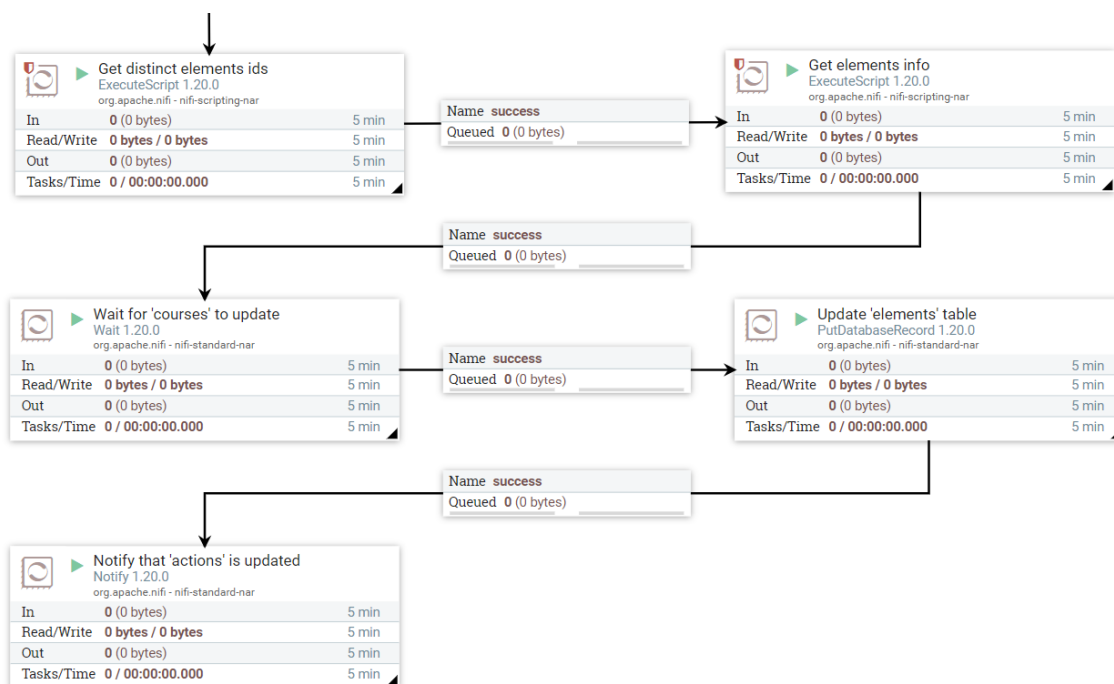


Рисунок 2. Фрагмент алгоритма миграции изменений элементов дисциплин в конечную базу данных

На основе Apache Superset реализована система для анализа образовательных данных. Apache Superset осуществляет визуализацию данных, получаемых при помощи SQL-запросов к базе данных DWH. Apache Superset позволяет пользователям создавать дашборды с графиками и отчётами, каждый из которых отображает определенные показатели в течение установленного периода времени (рисунок 3).

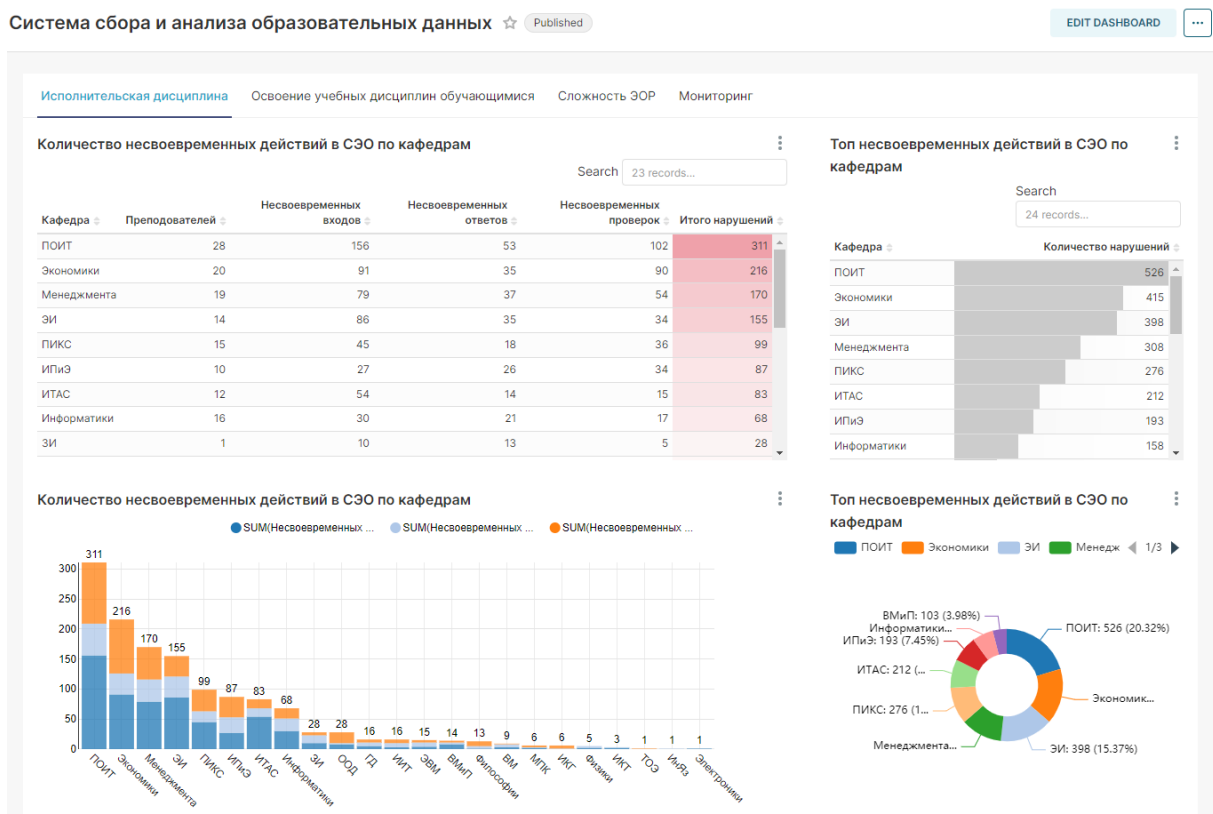


Рисунок 3. Дашборд системы мониторинга показателей образовательного процесса

На текущий момент реализован мониторинг следующих укрупнённых показателей образовательного процесса:

- динамика активности обучающихся в разрезе дисциплин;
- своевременность представления обучающимися контрольных заданий на проверку;
- своевременность проверки преподавателями выполненных обучающимися работ;
- обеспеченность образовательного процесса электронными методическими материалами, их тип, формы представления и актуализация;
- актуальность и полнота подключения пользователей к электронным кабинетам дисциплин и др.

Например, на рисунке 4 приведена тепловая карта обращений обучающихся к учебным материалам в разрезе дней недели и времени суток. Данная информация позволяет составить портрет среднестатистического обучающегося, сделать вывод о наиболее продуктивном времени обучения, а также выбрать оптимальное время для проведения технических работ.

На рисунке 5 приведен отчёт о выполнении тестов и практических работ обучающимися в разрезе учебных дисциплин. На основании данного отчёта работники, курирующие образовательный процесс, могут отследить отстающих студентов и принять меры по улучшению успеваемости. При соответствующей доработке системы возможно создание автоматизированной системы оповещения. Разработанная система может использоваться как полномасштабная система образовательного мониторинга, представляя данные о ходе образовательного процесса в разрезе отдельных обучающихся, учебных групп, факультетов, а также отдельных преподавателей, кафедр. Это позволит сократить затраты труда управленческого персонала и педагогических работников учреждений образования на осуществление образовательного мониторинга, повысить оперативность предоставляемой информации и предупредить появление возможных внештатных ситуаций в образовательном процессе, повысив качество образовательного процесса в целом.



Тепловая карта обращений к учебным материалам по дням недели

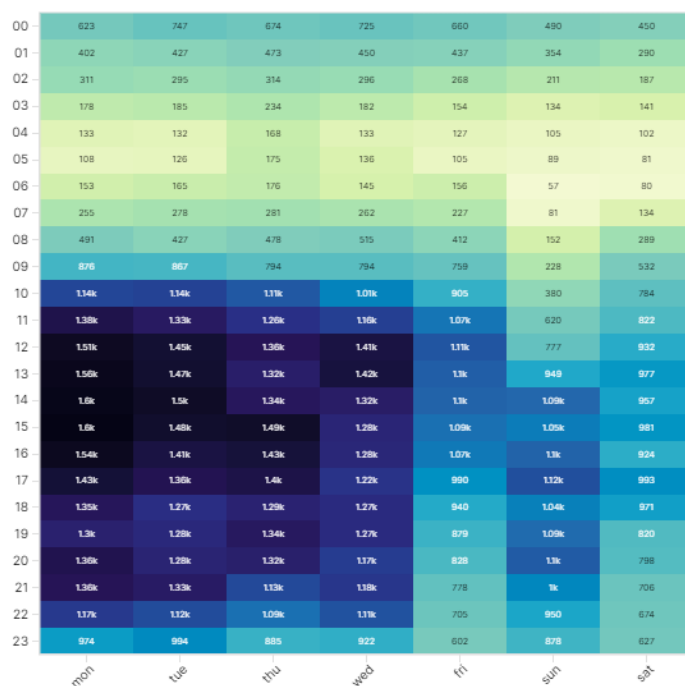


Рисунок 4. Тепловая карта обращений обучающихся к учебным материалам

Несвоевременная отправка студентами ДФПО тестов/работ на проверку

Show 200 entries Search 5703 records...

id события	ФИО студента	Группа	Дата оконч. уч. сем.	Название дисциплины	Заданий по программе	Заданий отправлено	Тестов по программе	Тестов пройдено	Последний вход
Владимировна			04						
17123	Софья Эвна	273951	2023-04-04	Английский язык. Часть 2	59	3	0	0	2023-04-18
17075	ич Полина	273951	2023-04-04	Основы компьютерной графики	3	3	2	2	2023-04-18
16945	Матвей Фролич	251051	2023-04-04	Основы алгоритмизации и программирования. Часть 2	5	3	5	1	2023-04-17
16937	Анна	251051	2023-04-04	Основы алгоритмизации и программирования. Часть 2	5	3	5	0	2023-04-17

1 2 3 4 5 6 7 ... 29

Рисунок 5. Отчёт о выполнении обучающимися тестов и практических работ

Особенностью системы является возможность гибкой настройки под конкретный образовательный процесс: добавление новых отчётов об интересующих показателях образовательного процесса, настройка дашбордов, разграничение доступа для групп пользователей.

Дальнейшие направления развития системы авторы видят в разработке системы оповещения субъектов образовательного процесса о необходимых действиях, а также построении индивидуальных траекторий и «жизненного цикла» обучающихся в системе электронного обучения на основе собранных данных.

### Список литературы

[1] Азаров, А.А. Совершенствование системы управления цифровой инфраструктурой университета: практика сетевого анализа / Азаров А.А., Бродовская Е.В., Лукушин В.А. // Высшее образование в России, Том. 32 №

2: научно-педагогический журнал, Россия, 2023 г./ отв.ред. Д.В. Давыдова. – Россия: Высшее образование в России, 2023. – С. 61–79.

[2] Крез, К. С. Вероятностные закономерности в цифровых следах пользователей системы электронного обучения учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» / Крез К. С., Дроздова М. А. // Электронные системы и технологии: материалы 59-й научной конференции аспирантов, магистрантов и студентов БГУИР, Минск, 17–21 апреля 2023 г. / редкол.: В.Ф. Алексеев [и др.]. – Минск : БГУИР, 2023. – С. 186–190.

[3] Раецкий, А. Д. Разработка отчета к системе MOODLE для организации контроля работы участников образовательного процесса / А. Д. Раецкий, С. А. Шлянин, Л. А. Ермакова // Теплотехника и информатика в образовании, науке и производстве : сборник докладов VI Всероссийской научно-практической конференции студентов, аспирантов и молодых учёных (ТИМ'2017) с международным участием (Екатеринбург, 11–12 мая 2017 г.). – Екатеринбург : УрФУ, 2017. – С. 244-248.

[4] Тербушева, Е.А. Аналитический потенциал платформы Moodle для мониторинга качества персонализированного обучения / Тербушева Е.А., Пиотровская К.Р. // Общество. Коммуникация. Образование, Том 12 № 4, Санкт-Петербургский, 2021 г. / редкол.: В.Е. Чернявская [и др.]. Санкт-Петербург : СПбПУ, 2021. – С. 19-34.

[5] Кустицкая, Т.А. Цифровой след из LMS Moodle для прогнозирования результатов обучения / Кустицкая Т.А., Есин Р.В., Носков М.В., Зыкова Т.В. // Информатизация образования и методика электронного обучения: цифровые технологии в образовании : материалы VI Международной научной конференции в трех частях, Том Часть 1, Красноярск, 20–23 сентября 2022 г. / редкол.: М. В. Носкова [и др.]. – Красноярск : КРПУ, 2022. – С. 90 – 94.

[6] Шнейдеров, Е. Н. Мониторинг количественных показателей образовательного процесса в системе электронного обучения на базе Moodle LMS / Шнейдеров Е. Н., Селиверстов Ф. Ф., Мигалевич С. А. // Дистанционное обучение – образовательная среда XXI века : материалы XII Международной научно-методической конференции, Минск, 26 мая 2022 г. / редкол.: Е. Н. Шнейдеров [и др.]. – Минск : БГУИР, 2022. – С. 98–100.

## **EDUCATIONAL PROCESS MONITORING SYSTEM FOR EDUCATIONAL PLATFORMS BASED ON MOODLE LMS**

***E.N. Schneiderov***

*Vice-Rector for Academic Affairs, Associate Professor of the Department of Information and Computer Systems Design of BSUIR*

***M.P. Batura***

*Doctor of Technical Sciences, Professor, Academician of the International Academy of Sciences of Higher Education*

***A. S. Tereshkova***

*Deputy Head of the Center for E-Learning Development of BSUIR*

***K.S. Krez***

*Master's student of BSUIR*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: shneiderov@bsuir.by*

**Abstract.** A system for monitoring indicators of the educational process for educational platforms based on Moodle LMS has been developed. The block diagram of the system is described, as well as the possibilities of its use on the example of the educational process in the BSUIR e-learning system of BSUIR.

**Keywords:** educational process monitoring, educational data analysis, Moodle LMS.

УДК 004.853

## ПОИСК ВИЗУАЛЬНО ПОДОБНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ



**М. М. Гресик**  
Студентка БГУИР КСиС  
mariya.gm.4@mail.ru



**С. Н. Нестеренков**  
Доцент кафедры программного  
обеспечения информационных  
технологий, Декан ФКСиС, Кандидат  
технических наук  
s.nesterenkov@bsuir.by

### **М. М. Гресик**

Студентка 4-ого курса специальности информатика и технологии программирования БГУИР. Проходит производственную практику в Центре информатизации и инновационных разработок в Белорусском государственном университете информатики и радиоэлектроники.

### **С. Н. Нестеренков.**

Кандидат технических наук, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры Программного обеспечения информационных технологий. Многократный автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

**Аннотация.** В данной статье рассматриваются вопросы поиска изображений по визуальному образцу. Предлагается методика разработки алгоритма оценки визуальной близости пары изображений на основе машинного обучения. Анализ и использование признаков изображения для построения специализированных поисковых индексов, используемых в алгоритме оценки визуальной близости, а также выбор таких признаков.

**Ключевые слова:** поиск изображений по визуальному образцу, машинное обучение, нейронные сети.

### **Введение.**

Развитие компьютерной техники и технологий привело к бурному росту количества цифровых изображений. Источников таких изображений множество — цифровые фотоаппараты и видеокамеры, видеорегистраторы, камеры видеонаблюдения, мобильные телефоны, сканеры и т. д. Появились цифровые архивы, содержащие миллионы снимков.

При таких объемах данных важное значение приобретает поиск нужных изображений.

Существует несколько направлений поиска по изображениям: поиск по содержанию (найти фотографию человека или изображение, к примеру, берёзы), поиск по визуальному образцу (найти изображения, похожие на заданное), поиск по описаниям (найти изображение, помеченное как «Совершенно секретно» или «Москва») и т. д.

Каждое из направлений поиска имеет свои особенности и сферы применения.

Целью работы является рассмотрение вопросов построения системы поиска по визуальному образцу в относительно больших (десятки тысяч) коллекциях изображений.

Поиск изображений по визуальному образцу является актуальной задачей, решению которой посвящено множество работ. Данный вид поиска заключается в извлечении существенных свойств изображений и построении на их основе сигнатур, используемых в дальнейшем для сравнения пар изображений. В каждую пару всегда входит изображение из коллекции и изображение-образец, подобное которому стремится найти пользователь.

Результатом сравнения является величина, называемая визуальным подобием

(тематической близостью) изображений.

С точки зрения информационного поиска данная оценка, выполненная человеком-экспертом, будет являться содержательной релевантностью, а рассчитанная в системе — релевантностью формальной.

Свойства оцениваются алгоритмами вычисления значений важных численных величин и называются признаками. Подобных признаков можно выделить достаточно много.

В общем случае они разделяются на две группы: глобальные и локальные. К глобальным признакам можно отнести основные цвета, текстуры, формы, значимые элементы всего изображения. Локальные признаки высчитываются для небольших частей (блоков) изображения. Для обозначения значений признаков также употребляют термин дескриптор. Набор признаков (вектор признаков), описывающий изображение для определенной задачи, называется сигнатурой.

Признаки, отобранные из множества альтернативных для их применения в некоторой задаче анализа изображений и вошедшие в сигнатуру, называют ключевыми. При поиске изображений по образцу анализируются не отдельные экземпляры, а пары изображений, которые сопоставляются друг другу.

Поэтому в отличие от других задач анализа в текущей проблеме от признаков изображений переходят к признакам пар, значения которых находятся как абсолютные разности значений соответствующих признаков каждого из изображений пары.

Такая несложная формула позволяет моментально их рассчитать для проиндексированных изображений и реализовать на практике поиск с приемлемыми временными характеристиками.

#### **Алгоритм определения визуального подобия изображений.**

Схематично предлагаемая методика работ для построения алгоритма определения визуального подобия изображений на основе машинного обучения показана на рисунке 1.

Ее можно представить в виде этапов, которые должны быть последовательно выполнены:

Этап 1. Из коллекции изображений формируется обучающая выборка. Она состоит из множества прецедентов, каждый из которых включает в себя пару изображений и степень их близости, оцененную экспертом.

Этап 2. Программно реализуются (при необходимости и разрабатываются) алгоритмы вычисления всевозможных признаков на изображениях.

Этап 3. Для каждой пары изображений вычисляются значения всех признаков.

Этап 4. Проводится факторный анализ признаков, в результате которого определяются главные признаки, значения которых будут использоваться в обучении нейронной сети и алгоритмы вычисления которых станут частью конечного алгоритма оценки визуального подобия изображений.

Этап 5. Обучается нейронная сеть. По завершению этого процесса запоминаются данные, характеризующие величины ошибок нейронной сети относительно входной выборки. Эти показатели преобразуются в полноту и точность — целевые значения оценки качества выполняемой работы.

Этап 6. Из коллекции изображений формируется тестовая выборка. Ее структура и характеристики (объем, соотношение близких и неблизких пар) должны соответствовать структуре и характеристикам обучающей выборки.

Этап 7. Вычисляются значения ключевых признаков всех изображений, входящих в тестовую выборку.

Этап 8. Прогоняется нейронная сеть на тестовой выборке с определением полноты и точности.

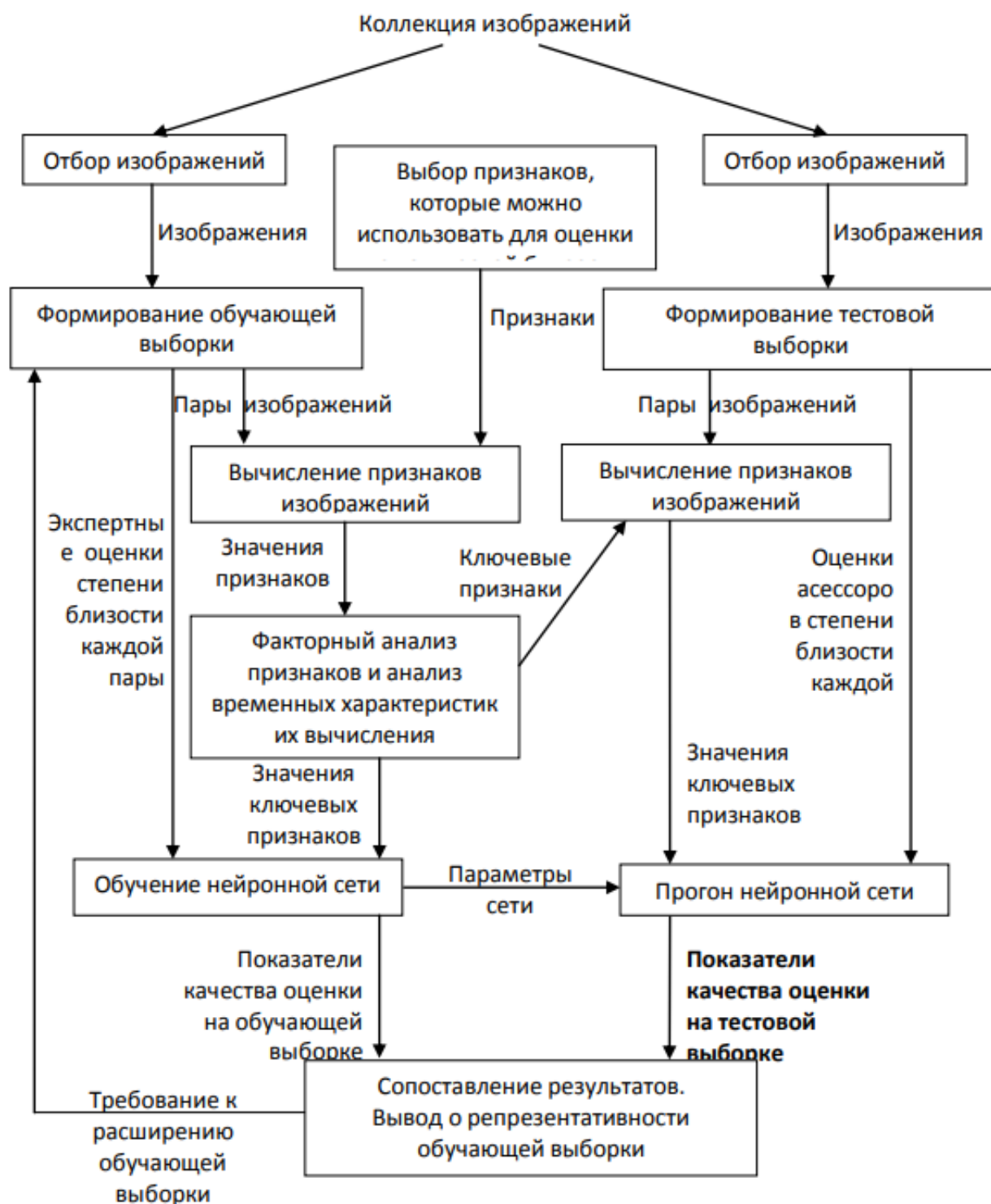


Рисунок 1. Схема построения алгоритма определения визуального подобия изображений

Этап 9. Сравниваются полнота и точность анализа нейронной сетью изображений обучающей выборки с полнотой и точностью анализа нейронной сетью изображений тестовой выборки.

При приблизительно одинаковых значениях соответствующих показателей работа считается завершенной.

### **Признаки изображений.**

В связи с тем, что поиск может осуществляться в больших коллекциях изображений, в качестве признаков мы рассматривали только те, которые можно просчитать для изображений заранее и занести в некое подобие поискового индекса (контрастность, грубость, направленность линейных образов, регулярности и шероховатости текстур).

По этой причине, нами не рассматривались те признаки, которые необходимо рассчитывать по парам изображений непосредственно при осуществлении поиска. Из 25 признаков на этапе выбора главных часть была отсеяна. Таким образом, к ключевым признакам были отнесены следующие:

- Средние значения компонент R, G, B изображения в цветовой модели RGB;
- Медиана яркости;
- Детализованность;
- Признак симметрии изображения;
- Признак наличия текстур;
- Пропорции сторон;
- Наличие лиц на изображении;
- Нормированные 16-ти уровневые гистограммы по компонентам Y (яркости), U, V изображения в цветовой модели YUV;
- Среднеквадратичное отклонение яркостей уменьшенных копий (размером 32x32 пикселя) образца и рассматриваемого изображения.

В качестве набора исходных данных использовалась коллекция Flickr [1].

В ней содержится 20000 фотографий разного качества и без единой темы; имеются фотоснимки людей, пейзажи, городские сцены и т.д. Фотографии сделаны при разном освещении, в помещениях, на улице и т.д.

Из данной коллекции в обучающую выборку были отобраны 500 пар, среди которых 125 являются тематически близкими, а 375 нет. Тестовой выборкой являлось специальное задание дорожки коллекции Flickr с результатами работы группы ассессоров.

Для формирования выборок и обучения нейронной сети применялся нейроимитатор Сигнейро [2], специализирующийся на задачах обработки и анализа изображений. Была сконструирована двухслойная нейронная сеть, структура которой приведена на рисунке 2.

Количество входов сети соответствует количеству ключевых признаков, выход сети — один. Для осуществления поиска изображений по визуальному подобию на основе полученной нейронной сети была разработана программная оболочка Axioma.

В процессе поиска она извлекает данные сигнатур (значения признаков) из поискового индекса, осуществляет вычисление меры близости образца к изображениям коллекции и ранжирует результаты поиска.

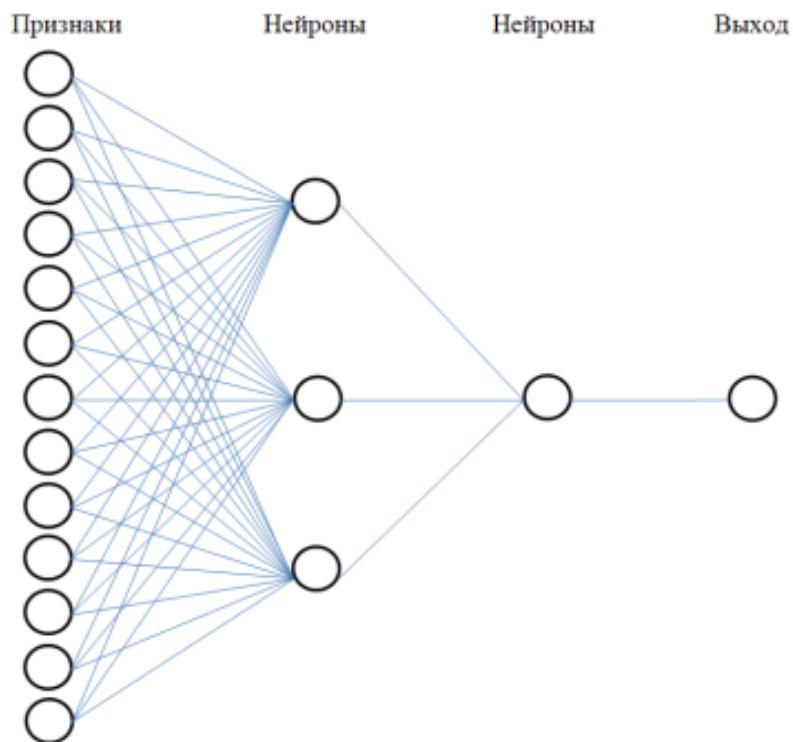


Рисунок 2. Структура двухслойной нейронной сети

График полнота-точность полученного нейросетевого алгоритма приведен на рисунке 3. На графике виден выступ, крайняя точка которого соответствует решению, полученному нейронной сетью.

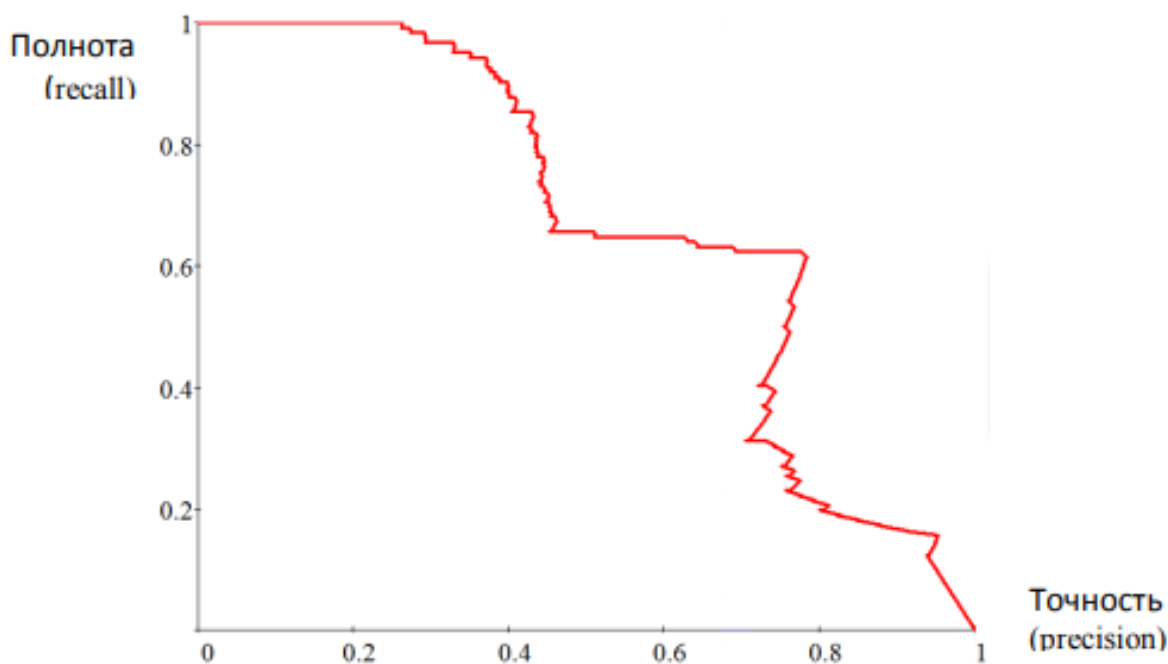


Рисунок 3. График полнота-точность

### **Заключение.**

Анализируя проведенную работу, были сделаны следующие выводы:

Полученные качественные характеристики построенного алгоритма имеют достаточно высокую оценку.

Вошедшая в оптимизационный критерий временная характеристика позволила минимизировать трудоемкость итогового алгоритма, который обладает линейным порядком временной сложности.

В дальнейшем, при увеличении объемов обучающей и тестовой выборок, качественные показатели могут быть улучшены или даже превзойдут значения аналогов, так как на данный момент имеются расхождения в показателях обучающей и тестовой выборок.

### **Список литературы**

[1] Коллекция изображений Flickr [Электронный ресурс]. – Электронные данные – Режим доступа: <http://romip.ru/ru/collections/flickr.html>. – Дата доступа: 24.03.2023

[2] Сигнейро. Нейронные сети в обработке и анализе изображений [Электронный ресурс]. – Электронные данные – Режим доступа: <http://www.signneuro.ru>. – Дата доступа: 24.03.2023

[3] Нестеренков, С.Н. Сравнительный анализ методов бинаризации изображений/ С.Н. Нестеренков, Д.В.Орлов, А. Н.Марков// Восьмая Международная научно-практическая конференция «BIG DATA and Advanced Analytics. BIG DATA и анализ высокого уровня», Минск, Республика, 11-12 мая 2022 год. - С. 104-109

[4] Нестеренков, С.Н. Использование генетического алгоритма для нахождения коэффициентов и структуры нейронной сети / С.Н. Нестеренков, К.П. Белов // Информационные технологии и системы 2018 (ИТС 2018) : материалы междунар. науч. конф., Минск, 25 окт. 2018 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2018. - С. 124-125.

[5] Нестеренков, С. Н. Модифицированный генетический алгоритм для обучения нейронной сети/ С. Н. Нестеренков, К. П. Белов // Информационные технологии и системы 2017 (ИТС 2017) : материалы междунар. науч. конф., Минск, 25 окт. 2017 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2017. - С. 204-205

## **SEARCH FOR VISUALLY SIMILAR IMAGES BASED ON MACHINE LEARNING**

***M. M. Hresik***  
*BSUIR KSiS student*

***S. N. Nesterenkov***  
*PhD Associate professor of department of  
the software of information technologies*

*Center for Informatization and Innovative Developments of the Belarusian State University of Informatics and Radio Electronics, Republic of Belarus  
Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: mariya.gm.4@mail.ru, s.nesterenkov@bsuir.by.*

**Abstract.** This article deals with the search for images by visual pattern. A technique for developing an algorithm for estimating the visual proximity of a pair of images based on machine learning is proposed. Image features that can be used in search algorithms are analyzed. A solution to the problem of choosing among them the key ones is presented.

**Keywords:** Image search by visual pattern, machine learning, neural networks.

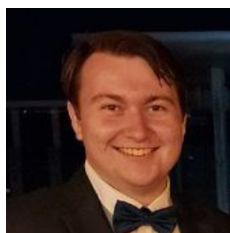


УДК 004.622

## BIG DATA: ПРОБЛЕМЫ И ТЕХНОЛОГИИ



**Д.И. Альховик**  
Студент 1 курса, кафедры  
информатики  
25350006@study.bsuir.by



**В.Д. Владымицев**  
Ассистент кафедры  
информатики, инженер  
программист ОИАСУ ЦИИР  
БГУИР  
v.vladymtsev@bsuir.by



**А.Н. Марков**  
Старший преподаватель,  
заместитель начальника  
Центра информатизации и  
инновационных разработок  
БГУИР  
a.n.markov@bsuir.by

**Д.И. Альховик**

Студент 1 курса специальности «Информатика и технологии программирования» БГУИР

**В.Д. Владымицев**

Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР

**А.Н. Марков**

Старший преподаватель кафедры ПОИТ, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Данный доклад рассматривает тему Big Data и его применение в различных областях, таких как медицина, финансы, транспорт, розничная торговля, государственное управление и интернет вещей. В докладе приведены конкретные примеры использования Big Data в каждой из этих областей, а также описаны потенциальные выгоды и преимущества, которые может принести использование данных большого объема. В заключении доклада подчеркивается, что примеры применения Big Data продолжают развиваться и расширяться, открывая новые возможности для улучшения бизнес-процессов и повышения эффективности в различных областях.

**Ключевые слова:** Big Data, анализ данных, медицина, финансы, транспорт, розничная торговля, государственное управление, IoT, примеры использования.

### Введение.

В современном мире мы сталкиваемся с огромными объемами данных, которые постоянно генерируются из различных источников. Эти данные могут содержать ценную информацию, но для их обработки и анализа необходимы специальные технологии и инструменты. В связи с этим, концепция Big Data становится все более актуальной и востребованной.

Однако, работа с Big Data также включает в себя ряд проблем и вызовов, связанных с объемом, скоростью, разнообразием и качеством данных, а также безопасностью и доступностью информации. Необходимость обработки больших объемов данных ставит перед нами задачу разработки новых технологий и инструментов, которые позволят эффективно работать с Big Data.

В данном докладе мы рассмотрим основные проблемы и вызовы, связанные с Big Data, а также технологии и инструменты, которые используются для работы с такими данными. Мы также рассмотрим примеры применения Big Data в различных областях, таких как медицина, банковская сфера, ритейл, транспорт и наука. В конце доклада мы обсудим значение Big Data в будущем и дадим рекомендации для дальнейших исследований.

### Проблемы Big Data.

Big Data представляет собой набор данных, которые могут быть слишком большими, сложными и разнообразными для обработки традиционными методами. Объем и скорость

создания данных постоянно растет, что приводит к ряду проблем, связанных с их обработкой и анализом. Ниже мы рассмотрим некоторые из основных проблем, связанных с Big Data:

1. Объем данных: с постоянным ростом количества данных, с которыми мы работаем, существует необходимость в большем объеме вычислительных ресурсов и хранилищ для их сохранения и обработки.

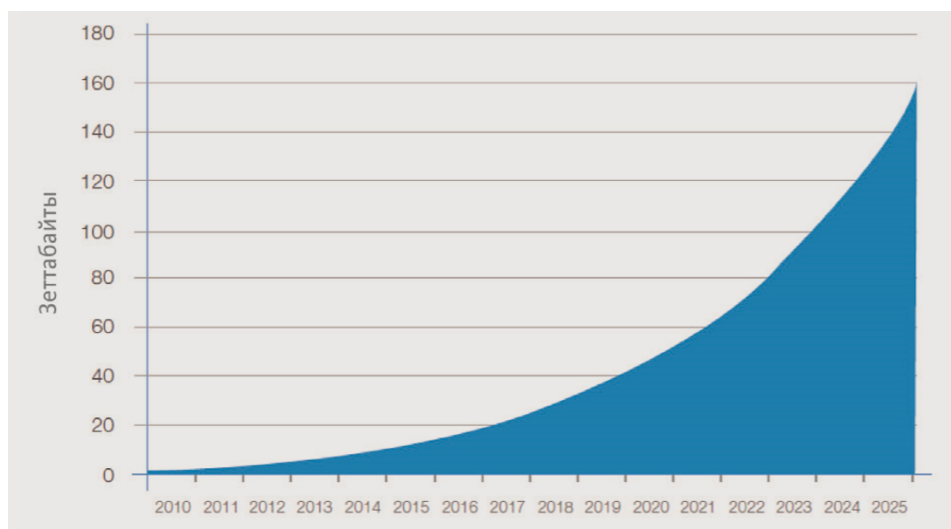


Рисунок 1 – Рост объема данных в мире.

2. Скорость обработки данных: некоторые приложения, такие как интернет-поиск или реклама в режиме реального времени, требуют мгновенной обработки и анализа данных. Это ставит перед нами задачу оптимизации производительности и скорости обработки данных.

3. Разнообразие данных: данные могут поступать из различных источников и в различных форматах. Это усложняет процесс их обработки и анализа, так как необходимы различные инструменты и методы для работы с различными типами данных.

4. Качество данных: не все данные являются полезными или точными. Некоторые могут быть неполными, некорректными или содержать ошибки. Это может привести к неточным выводам и решениям, основанным на этих данных.

5. Безопасность данных: с увеличением объема данных также возрастает риск их утечки, взлома или кражи. Это может привести к серьезным последствиям для компаний и организаций, которые хранят и обрабатывают конфиденциальные данные.

6. Недостаток опытных специалистов: работа с Big Data требует специализированных знаний и навыков. Однако, на данный момент спрос на таких специалистов значительно превышает их количество, что создает дополнительные проблемы в работе с Big Data.

### **Технологии обработки Big Data.**

Технологии обработки Big Data развиваются с высокой скоростью, чтобы обеспечить эффективную и быструю обработку большого объема данных. В данном контексте, стоит отметить такие технологии как:

Apache Hadoop [1]: распределенная система хранения и обработки данных, которая использует большое количество узлов серверов для распределения задач по обработке данных. Она позволяет обрабатывать большие объемы данных параллельно и снижает время обработки.

Apache Spark [7]: открытая платформа обработки больших данных, которая позволяет выполнять вычисления в памяти, что ускоряет обработку данных и улучшает производительность. Она также поддерживает обработку потоков данных, машинное обучение и графовые вычисления.

NoSQL базы данных [4]: это альтернатива реляционным базам данных, которая может обрабатывать неструктурированные данные и гораздо лучше масштабируется для работы с большими объемами данных. Примеры NoSQL баз данных включают MongoDB и Cassandra.

HBase [1]: это распределенная база данных, которая предназначена для работы с большими объемами структурированных данных. Она позволяет хранить, обновлять и извлекать данные в режиме реального времени, а также обеспечивает масштабируемость и отказоустойчивость.

Apache Flink [4]: это открытая платформа обработки потоков данных и батч-процессинга. Она позволяет обрабатывать данные в реальном времени, обеспечивает высокую производительность и масштабируемость.

Эти технологии являются только некоторыми из инструментов, которые используются для обработки больших объемов данных, и их выбор зависит от конкретных требований и задач, которые необходимо решить.

### **Примеры использования Big Data.**

Примеры использования Big Data разнообразны и могут включать в себя различные отрасли и области. Некоторые примеры использования Big Data включают:

1. Медицина: использование данных о здоровье, пациентах и их истории болезни для улучшения диагностики и лечения, а также оптимизации системы здравоохранения.

2. Финансы: анализ данных о транзакциях, потребительском поведении и финансовых рынках для прогнозирования трендов и управления рисками.

3. Транспорт: использование данных о движении транспорта и поведении водителей для улучшения безопасности дорожного движения и оптимизации транспортных потоков.

4. Розничная торговля: анализ данных о покупках, поведении потребителей и тенденциях рынка для улучшения продуктового ассортимента, ценообразования и управления запасами.

5. Государственное управление: использование данных о социально-экономическом развитии, демографических трендах и других параметрах для принятия решений в области политики и управления ресурсами.

6. Интернет вещей (IoT): использование данных, собираемых от различных устройств и датчиков, для управления умными домами, городами и промышленными системами [8-13].

Примеры использования Big Data продолжают развиваться и расширяться, открывая новые возможности для улучшения бизнес-процессов и повышения эффективности в различных областях.

### **Заключение.**

Big Data является важной технологией в современном мире, позволяющей организациям собирать, хранить и анализировать большие объемы данных. Примеры использования Big Data разнообразны и могут быть найдены в различных отраслях и областях, включая медицину, финансы, транспорт, розничную торговлю, государственное управление и IoT. Big Data помогает организациям принимать более обоснованные решения, повышать эффективность бизнес-процессов и улучшать качество продуктов и услуг. Однако, использование Big Data также представляет вызовы, такие как необходимость обеспечения безопасности и конфиденциальности данных, а также необходимость наличия высококвалифицированных специалистов для работы с данными. Несмотря на эти вызовы, использование Big Data остается важной стратегией для организаций, желающих оставаться конкурентоспособными в современном цифровом мире.

### **Список литературы**

- [1] IBM. "Big data analytics." IBM. <https://www.ibm.com/analytics/big-data-analytics>  
[2] Amazon. "Big data." Amazon Web Services. <https://aws.amazon.com/big-data/> [3] Google. "What is big data?" Google Cloud. <https://cloud.google.com/big-data/what-is-big-data>  
[4] C. Mohan. "Big data: Promises, challenges and opportunities." IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 6, 2014, pp. 1526-1537.

- [5] V. G. Cerf and R. E. Kahn. "A protocol for packet network intercommunication." IEEE Transactions on Communications, vol. 22, no. 5, 1974, pp. 637-648.
- [6] D. L. Mills. "Internet time synchronization: the Network Time Protocol." IEEE Transactions on Communications, vol. 39, no. 10, 1991, pp. 1482-1493.
- [7] M. Zaharia, et al. "Apache Spark: a unified engine for big data processing." Communications of the ACM, vol. 59, no. 11, 2016, pp. 56-65.
- [8] McKinsey & Company. "The age of analytics: Competing in a data-driven world." McKinsey & Company, December 2016.
- [9] Forbes. "10 powerful examples of big data analytics in healthcare."
- [10] IBM. "Big data in banking: How to use big data analytics."
- [11] KD Nuggets. "7 innovative examples of big data in action."
- [12] Gartner. "Gartner Top 10 Strategic Technology Trends for 2022."
- [13] SAS. "Big data analytics." SAS

## **BIG DATA: PROBLEMS AND TECHNOLOGIES**

***D. I. Alkhovik***

*Student of BSUIR, software  
engineer*

***V.D. Vladymtsev***

*Bechalar engineer system-  
programmer, Department of  
Integrated Automated Control  
Systems*

***A.N. Markov***

*Senior lecturer of the department,  
Deputy head of the Center for  
Informatization and Innovative  
Developments*

*Center for Informatization and Development of the Belarusian University of State Informatics and  
Radioelectronics, Republic of Belarus*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*

*E-mail: 25350006@study.bsuir.by, v.vladymtsev@bsuir.by, a.n.markov@bsuir.by*

**Annotation.** This report examines the topic of Big Data and its application in various fields, such as medicine, finance, transport, retail, public administration and the Internet of Things. The report provides specific examples of the use of Big Data in each of these areas, as well as describes the potential benefits and advantages that the use of large-volume data can bring. In conclusion, the report emphasizes that the examples of Big Data applications continue to develop and expand, opening up new opportunities for improving business processes and increasing efficiency in various areas.

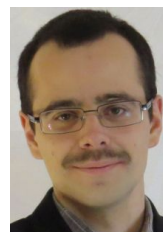
**Keywords:** Big Data, data analysis, medicine, finance, transport, retail, public administration, IoT, usage examples.

УДК 004.67

## ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ ДЛЯ РАННЕГО ОБНАРУЖЕНИЯ ЛЕСНЫХ ПОЖАРОВ



**Д.Д. Дронов**  
магистрант гр. 155841,  
кафедра ЭВМ, БГУИР



**Д.Ю. Перцев**  
кандидат технических наук,  
кафедра ЭВМ, БГУИР

### **Д.Д. Дронов**

Является магистрантом Белорусского государственного университета информатики и радиоэлектроники по специальности 1-40 80 01 «Компьютерная инженерия (Хранение и обработка данных)».

### **Д.Ю. Перцев**

Окончил аспирантуру Белорусского государственного университета информатики и радиоэлектроники (2016), защитил диссертацию в совете 05.13.01 (2020). Является доцентом кафедры ЭВМ.

**Аннотация.** Выявлены факторы для раннего обнаружения лесных пожаров, определены методы анализа и выполнен многофакторный анализ данных, представлена визуализация результатов анализа данных.

**Ключевые слова:** Интеллектуальный анализ данных, Многофакторный анализ, Регрессионный анализ, Визуализация многомерных данных, Метод главных компонент, Лесные пожары.

### **Введение**

Лесные пожары – это серьезная угроза для окружающей среды и человеческой жизни. Они могут привести к уничтожению животных и растительных видов, земельных угодий и инфраструктуры. Пожары могут иметь серьезные последствия для экосистем лесов, угрожая их биоразнообразию и стабильности [1].

Необходимо принимать меры для предотвращения пожаров и контроля их распространения, а также обеспечивать эффективное и быстрое реагирование на возможные пожары для минимизации их негативных последствий.

В настоящее время для обнаружения лесных пожаров используются методы, пригодные для различных условий и территорий. Наиболее распространёнными методами являются:

- камеры видеонаблюдения. Они располагаются на высоте кромок деревьев, имеют обзор 360 градусов и способны обнаружить дым и его приблизительное местоположение. Возможны ложные обнаружения из-за естественных и техногенных испарений, которые похожи на дым;
- наблюдение со спутников. Получение снимков со спутников может занимать продолжительное время – от 2 до 12 часов;
- наблюдение с вышек и патрулирование;
- сообщения от граждан;
- беспилотные летательные аппараты.

### **Описание системы**

Система раннего обнаружения лесных пожаров представляет собой множество устройств, объединённых в сеть ячеистой топологии [2][3] по протоколам IoT (Internet of Things). Каждое устройство содержит датчик температуры, углекислого газа, водорода и углеводородных газов [4][5][6].

В данной статье рассматривается метод многофакторного анализа данных для обработки информации от датчиков и других источников, которые собирают данные о пожаре в лесу. Все данные собираются в таблицу, пример которой представлен в таблице 1.

Входные данные включают дату и время, идентификатор устройства, геопозицию и показания с датчиков. Дополнительные факторы могут включать информацию о количестве осадков, силе и направлении ветра от метеорологических служб. Все эти данные обрабатываются с помощью многофакторного анализа данных.

Таблица 1. Пример данных, получаемых от устройств

#	ID	Дата	Время	GPS-широта	GPS-долгота	Температура	CO <sub>2</sub>	H <sub>2</sub>	C <sub>x</sub> H <sub>x</sub>
1	1	05.06.2022	15:00	29	52	28	0,0143	0,2906	0,9413
2	6	05.06.2022	15:01	34	47	28	0,0075	0,2603	1,4221
3	9	05.06.2022	15:02	37	44	28	0,0074	0,2139	0,9248
4	4	05.06.2022	15:03	32	49	28	0,0076	0,2579	1,9364
5	8	05.06.2022	15:04	36	45	28	0,0128	0,2607	0,1835
6	2	05.06.2022	15:05	30	51	28	0,0164	0,2237	0,0473
7	7	05.06.2022	15:06	35	46	28	0,0092	0,2095	0,3126
8	5	05.06.2022	15:07	33	48	28	0,0010	0,2297	0,6582
9	7	05.06.2022	15:08	35	46	28	0,0180	0,2136	1,8317
10	3	05.06.2022	15:09	31	50	28	0,0051	0,2329	1,4122

### Многофакторный анализ данных

Метод многофакторного анализа данных является мощным инструментом, который позволяет определить взаимосвязи между несколькими переменными и определить значимые факторы, влияющие на исходную переменную.

Для данной системы анализ данных начинается с предварительной обработки входных данных, которые очищаются от выбросов, пропущенных значений и ошибок измерения. Далее выполняется нормализация данных, что позволяет сравнить различные переменные на основе их значений.

После этого строится модель многофакторного анализа данных. Данная модель включает в себя элементы, которые могут оказывать влияние на исходную переменную, – наличие пожара в лесу. Каждый фактор оценивается на основе коэффициента корреляции с исходной переменной [7].

Было обнаружено, что наиболее сильная связь наблюдается между температурой и содержанием углеродных газов. Кроме того, было обнаружено, что сила и направление ветра существенно влияют на распространение пожара.

Результаты анализа показали, что данные можно разделить на несколько групп, основанных на геопозиции и количестве осадков. Таким образом, можно определить области, наиболее подверженные пожароопасности в зависимости от географического расположения и количества осадков.

### Использование метода главных компонент (PCA)

Одним из методов многофакторного анализа данных является метод главных компонент (PCA). Он позволяет уменьшить размерность данных, выявить скрытые зависимости между факторами и сократить количество используемых переменных.

Представим, что мы имеем набор данных, включающий множество различных показателей, таких как температура, количество осадков, влажность, направление и скорость

ветра. Каждый показатель можно рассматривать как переменную. Такой набор данных может быть представлен в виде матрицы, где каждый столбец соответствует одному из показателей, а каждая строка - конкретному наблюдению.

Метод главных компонент позволяет определить, какие из показателей наиболее важны и объясняют наибольшую долю изменчивости в данных. Эти показатели называются главными компонентами. Остальные показатели могут быть отброшены или использованы в более узкой области исследования.

Таким образом, включение данных о метеорологических условиях в систему раннего обнаружения лесных пожаров может значительно улучшить ее точность и помочь предотвратить развитие пожаров в будущем.

### **Регрессионный анализ**

Одним из методов анализа данных, который может быть применен для системы раннего обнаружения лесных пожаров, является регрессионный анализ, который позволяет определить связь между зависимой переменной и набором независимых, и использовать эту связь для прогнозирования значений.

В случае системы раннего обнаружения лесных пожаров, зависимой переменной может быть вероятность возникновения пожара, а независимыми переменными – температура, содержание углекислого газа, концентрация водорода и других газов, а также осадки, направление и скорость ветра. Для построения регрессионной модели необходимо использовать исторические данные о значениях этих переменных в периоды, когда происходили пожары.

Для построения регрессионной модели необходимо выбрать подходящий тип модели, определить зависимую и набор независимых переменных, а также оценить коэффициенты регрессии и проверить значимость полученных результатов.

После построения модели можно использовать ее для предсказания вероятности возникновения пожара на основе текущих значений переменных. Для этого необходимо ввести значения независимых переменных в модель и получить прогноз для зависимой. Если прогнозируемая вероятность возникновения пожара выше заданного порогового значения, система может автоматически сигнализировать об этом и предпринимать необходимые меры.

Регрессионный анализ может быть полезным инструментом для системы раннего обнаружения лесных пожаров, позволяя на основе исторических данных о пожарах и погодных условиях определить вероятность возникновения новых пожаров и своевременно предпринять меры для их предотвращения.

### **Визуализация данных**

Для наглядного представления собранных данных используется геоинформационная система [8]. В зависимости от настройки параметров, возможно отображение датчиков и информации с них, либо раскрашивание карты в соответствии с риском возникновения пожара.

Лесной пожар вызывает отклонения значений датчиков. При этом различные виды пожара проявляются по-разному. Поэтому, при выявлении критериев важно опираться, по возможности, на данные с нескольких устройств, соседствующих друг с другом.

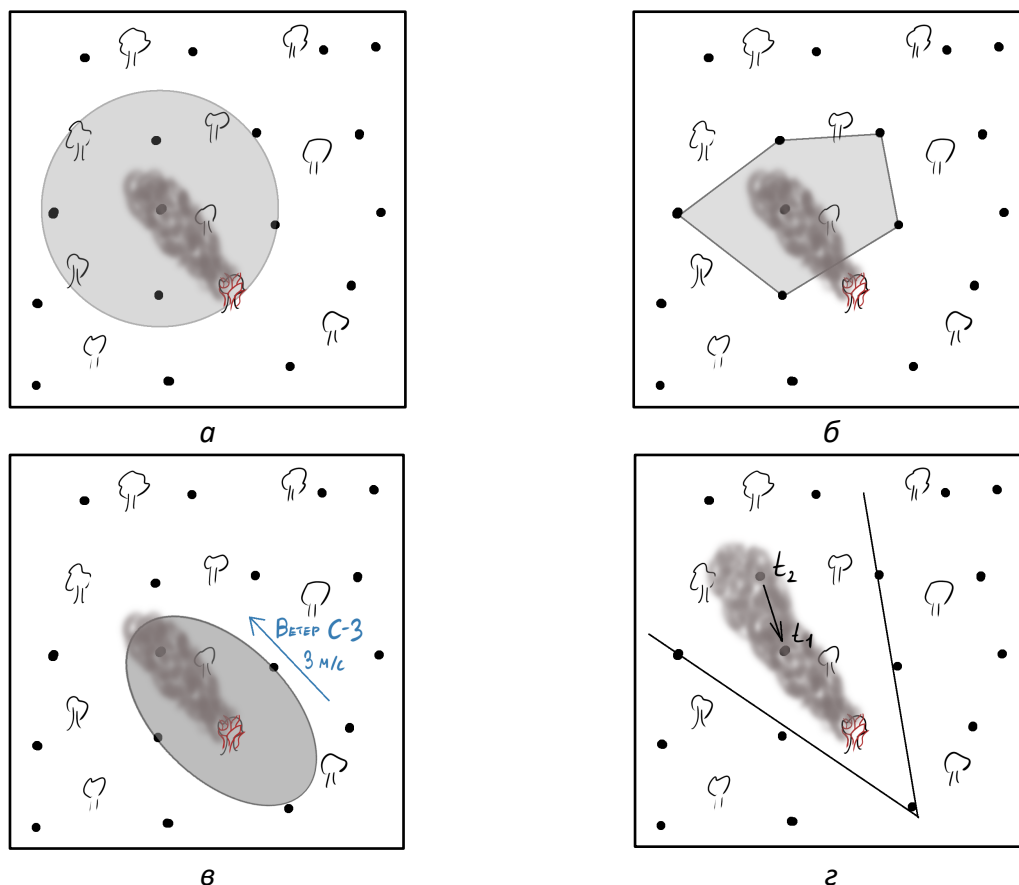


Рисунок 1. Определение области поражения: а - одно устройство; б – полигональное; в – с учётом метеоданных; г – на основе нескольких датчиков

Важно отметить, что в случае выявления признаков лесного пожара одним устройством, возможная область поражения описывается в общем случае радиусом, подходящим для мониторинга (рисунок 1.а), либо полигоном, границы которого определяются расстоянием между соседними датчиками (рисунок 1.б). Чтобы уточнить область поражения возможно использовать метеоданные, такие как скорость и направление ветра. В таком случае, возможная область поражения будет смещена в направлении, противоположном направлению ветра (рисунок 1.в). Если же один из датчиков обнаружил признаки пожара в момент времени  $t_1$ , а второй – в момент времени  $t_2$ , можно сделать вывод о направлении распространения продуктов горения (рисунок 1.г).

### Пример анализа данных

В качестве примера анализа данных рассмотрим определение факторов, влияющих на вероятность возникновения лесного пожара в Слонимском районе Гродненской области Республики Беларусь. Для этого были собраны данные о погодных условиях, количестве осадков на территории лесного массива и возникновении пожаров. Данные были обработаны и проанализированы, чтобы выделить наиболее значимые факторы, влияющие на вероятность возникновения лесного пожара. В результате анализа данных методом PCA было выявлено, что наибольшее влияние на вероятность возникновения пожара оказывают три фактора: температура, концентрация углекислого газа и скорость ветра. Более высокие значения этих параметров соответствуют более высокой вероятности возникновения пожара. В результате анализа с помощью авторегрессионной модели ARIMAX [9] удалось построить модель риска возникновения лесного пожара, которая учитывает сезонные изменения исследуемых факторов.



Таким образом, были выделены ключевые факторы, влияющие на вероятность возникновения пожара в лесном массиве и построена модель для своевременного обнаружения.

### **Заключение**

В данной работе была представлена система раннего обнаружения лесных пожаров на основе IoT устройств, объединенных в сеть ячеистой топологии. Были рассмотрены факторы, влияющие на пожароопасность, и выполнен анализ данных с помощью многофакторного анализа. Система раннего обнаружения лесных пожаров является важной мерой для предотвращения возникновения и быстрой локализации пожаров. Для ее эффективности необходимо учитывать множество различных факторов, таких как расположение устройств, типы датчиков и метеорологические условия.

### **Список литературы**

- [1] Пожары меняют климат, даже когда мы об этом забываем [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://greenpeace.ru/blogs/2019/09/23/pozhary-menjajut-klimat-dazhe-kogda-my-ob-jetom-zabyvaem/> Дата доступа: 15.03.2023
- [2] Wireless Mesh Networks in IoT Networks / Yu Liu, Xiangdong Qiu [et al.] // 2017 International Workshop on Electromagnetics: Applications and Student Innovation Competition, 2017. – С.55-57.
- [3] Early Forest Fire Detection Using a Protocol for Energy-Efficient Clustering with Weighted-Based Optimization in Wireless Sensor Networks / Puneet Kaur [et al.] // Applied sciences, 2023. – <https://doi.org/10.3390/app13053048>.
- [4] Early forest fire detection using low-energy hydrogen sensors / K. Nörthemann [et al.] // Journal of Sensors and Sensor Systems, 2013. – С.171-177.
- [5] Wildland fire emissions, carbon, and climate: Emission factors / Shawn Urbanski // Forest Ecology and Management, 2014. – С.51-60.
- [6] Early forest fire detection and verification using optical smoke, gas and microwave sensors / Wolfgang Krüll [et al.] // Procedia Engineering, 2012. – С.584-594.
- [7] «Сырые» данные и некоторые рецепты их «приготовления» / М.М. Татур [и др.] // Информационные системы и технологии = Information Systems and Technologies : материалы междунар. науч. конгресса по информатике. В 3 ч. Ч. 1, 2022. – С.194-203.
- [8] Зиновьев, А.Ю. Визуализация многомерных данных / А.Ю. Зиновьев. – М. : Изд-во КГТУ, 2000. – 180 с.
- [9] Тихонов, Э.Е. Методы прогнозирования в условиях рынка: учебное пособие / Э.Е. Тихонов – Невинномысск, 2006. - 221 с.

## **DATA MINING FOR EARLY DETECTION OF FOREST FIRES**

**D. Dronov**

*master's student of group 155841,  
Electronic Computing Machines Department,  
BSUIR*

**D. Pertsau**

*PhD,  
Electronic Computing Machines  
Department, BSUIR*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*

*E-mail: [bsuir.dmitriy.dronov@gmail.com](mailto:bsuir.dmitriy.dronov@gmail.com), [pertsev@bsuir.by](mailto:pertsev@bsuir.by)*

**Abstract.** Factors for early detection of forest fires are identified, analysis methods are determined and multivariate data analysis is performed, visualization of data analysis results is presented.

**Key words:** Data Mining, Multifactor analysis, Regression analysis, Visualization of multidimensional data, Principal component analysis, Forest fires.

УДК 004.021:004.75

## ИГРА ДОПОЛНЕННОЙ РЕАЛЬНОСТИ «БИТВА С РОБОТАМИ»



**У.П. Курилина**

Учащаяся Учреждения образования «Национальный детский технопарк», учащаяся ГУО «Гимназия г. Добруша»



**Д.А. Филон**

Учащийся Учреждения образования «Национальный детский технопарк», учащийся ГУО «Речицкий районный лицей»



**М.С. Ильясова**

магистрант кафедры инженерной психологии и эргономики БГУИР



**Ф.В. Усенко**

магистрант кафедры инженерной психологии и эргономики БГУИР



**Л.Р. Коркин**

магистр технических наук, ассистент кафедры инженерной психологии и эргономики БГУИР



**А.М. Прудник**

доцент кафедры инженерной психологии и эргономики БГУИР, кандидат технических наук, доцент [aleksander.prudnik@bsuir.by](mailto:aleksander.prudnik@bsuir.by)

### **У.П. Курилина**

Обучается в гимназии г. Добруша. Область научных интересов связана с технологиями виртуальной и дополненной реальности.

### **Д.А. Филон**

Обучается в Речицком районном лицее. Область научных интересов связана с разработкой приложений виртуальной и дополненной реальности.

### **М.С. Ильясова**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с автоматизированным тестированием информационных систем.

### **Ф.В. Усенко**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой манипуляторов для дистанционного управления.

### **Л.Р. Коркин**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с системами распознавания снимков.

### **А.М. Прудник**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с взаимодействием человека с компьютером, интерфейсами информационных систем, пользовательскими интерфейсами, *front-end web development*.

**Аннотация.** Рассматриваются компоненты, объекты и структуры, необходимые для создания AR-приложения на базе платформы Unity. Также рассматриваются некоторые аспекты разработки приложения

на реальном проекте «Битва с роботами», написанном в процессе работы. Помимо этого, в работе были применены методы тестирования проекта во время разработки и после окончательной сборки.

**Ключевые слова:** AR-технологии, виртуальная реальность, Unity, C#, Git, Blender, Visual Studio

### **Введение.**

В настоящий момент индустрия игр расширенной (XR), т.е. виртуальной, дополненной и смешанной реальности (VR, AR, MR) динамично развивается. Предполагается, что в 2023 г. оборот достигнет USD 31,12 млрд [1], а число активных пользователей — 1,7 млрд в 2024 г.

Основными ограничениями на дальнейшее внедрение и использование технологий XR являются (в порядке убывания): громоздкая гарнитура и технические сбои при её эксплуатации; отсутствие или недостаток контента; нежелание бизнеса, ориентированного на конечного потребителя, заниматься разработкой направления XR; правовые риски; финансирование инвестиций; стоимость для потребителей. Тем не менее, в игровой индустрии AR имеет большой потенциал.

Рынок также демонстрирует сильный тренд на разработку мобильных игр. Например, общий оборот рынка мобильных игр в 2021 г. составил USD 93,2 млрд [3], увеличившись на 7,3% по сравнению с предыдущим годом.

Целью проекта является разработка приложения AR «Битва с роботами» для демонстрации её возможностей дополненной реальности посредством игры, в которой будут отражены основные возможности дополненной реальности.

Основные преимущества игр AR на мобильных устройствах или в очках AR по сравнению со стандартными мобильными играми:

- физическая активность — возможность играть в игру, являющейся здоровой физической нагрузкой;
- игровой процесс без помощи рук — возможность играть в игру, в которой руки свободны от удержания устройства;
- реальный опыт от первого лица — по сравнению с мобильными устройствами, в очках AR это было похоже на настоящий опыт от первого лица;
- социальное взаимодействие лицом к лицу — в отличие от мобильных игр, в играх AR во время игрового процесса можно взаимодействовать и общаться с другими реальными игроками, наблюдая за их реакцией и общаясь с ними естественным образом;
- аутентичные элементы управления человеческим телом — игра AR в очках AR позволяет использовать жесты тела, недоступные в стандартных мобильных играх

Для достижения поставленной цели необходимо было решить следующие задачи:

1. Провести обзор аналогов разрабатываемой системы и выполнить сравнительный анализ их достоинств и ограничений;
2. Проанализировать технологии, подходящие для создания мобильного приложения с использованием дополненной реальности;
3. Спроектировать и реализовать приложение, используя эффективные шаблоны проектирования и разработки мобильных приложений;
4. Провести тестирование мобильного приложения.

Актуальность выполненного проекта заключается в предоставлении картины искусственной среды, разработанной с помощью программного решения, в котором доминируют зрение и звук, благодаря чему достигается максимальный эффект погружения в неё.

Практическая значимость тематики проектов заключается в том, что были получены готовые программные продукты, взаимодействуя с которыми пользователь может получить новый опыт, отработать определенные навыки, а также проводить время физически активно.

### Основная часть.

Для реализации приложения был выбран поджанр однопользовательского шутера от первого лица (FPS), что предполагает восприятие участником игрового процесса глазами главного героя.

В рамках данной игры пользователь сражается с противниками, представленными в виде виртуальных роботов. Враги появляются на горизонтальных поверхностях, отображаемых на экране устройства, и перемещаются к игроку, используя комбинацию стрельбы и передвижения в его направлении.

Задача участника игры заключается в уничтожении противников.

К основным функциям приложения относятся возможность воспроизведения выстрелов, настройка звука на экране главного меню, просмотр истории персонажа, за которого пользователь играет, отображение шкалы здоровья как у игрока, так и у противников. Если в процессе игры вражеская сторона поражает игрока, то его уровень здоровья уменьшается, аналогично, при поражении врага игрок набирает очки. Когда шкала здоровья опускается до нуля, то на экране появляется соответствующее уведомление о гибели персонажа и проигрывается анимация, эмулирующая данный процесс.

Работы велись с использованием кроссплатформенной среды разработки компьютерных игр Unity, языка программирования C# и программного обеспечения для создания трёхмерной графики Blender. Также в процессе работы использовалась распределённая система управления версиями Git.

В проекте, разрабатываемом в среде Unity, использовалась система сцен, которые представляли собой отдельные файлы, содержащие игровые миры с их объектами, сценариями и настройками. Каждому объекту присваивалось свое уникальное название, тег и слой, определяющие его отображение и взаимодействие в игре. Кроме того, для каждого объекта на сцене был определён компонент *Transform*, который хранил информацию о его местоположении, ориентации и размерах в пространстве игры по трём осям. Так, например, скрипт, представленный на рисунке один, отвечает за поворот противника к игроку.

```
public GameObject Camera;
// Update is called once per frame
void Update()
{
    transform.LookAt(Camera.transform, Vector3.up);
    transform.rotation = new Quaternion(0.0f, transform.rotation.y, 0.0f,
transform.rotation.w);
}
```

Рисунок 1. Скрипт, отвечающий за поворот противника к игроку

Кроме того, объектам с видимой геометрией задавался компонент *Mesh Renderer*, обеспечивающий видимость модели объекта. Также для объектов применялись коллайдеры, позволяющие взаимодействовать объектам посредством коллизий.

Для реализации механики стрельбы в проект был добавлен метод *Raycast*, используемый в скрипте, отображенном на рисунке 2.

Он представляет собой виртуальный луч, исходящий из определенной точки объекта в заданном направлении с определенной длиной (возможно, бесконечной), для обнаружения коллизий с другими объектами. Этот метод позволяет «выстрелить» из объекта и проверить попадание в коллайдер.

```
public void Fire()
{
    Ray ray = new Ray(Cam.transform.position, Cam.transform.forward);
    Debug.DrawRay(Cam.transform.position, Cam.transform.forward*100f,
Color.yellow);
    RaycastHit hit;
    if (Physics.Raycast(ray, out hit))
    {
        GameObject obj = hit.collider.gameObject;
        obj.GetComponent<HealthBarEnemy>().Damage(20f);
    }
}
```

Рисунок 2. Скрипт, отвечающий за выстрел игрока и нанесение урона в случае попадания

Если произошло попадание в противника или игрока, то его шкала здоровья уменьшается. Для реализации шкалы здоровья (в дальнейшем — *HealthBar*) были приняты следующие шаги: создан родительский объект, который представляет собой полосу здоровья.

Затем, каждый противник наследует эту полосу и создает копию в его дочерних объектах. Управление *HealthBar* осуществляется посредством скриптов, один из которых представлен на рисунке 3.

```
public void Damage(float DamageHP)
{
    HP -= DamageHP;
    HelthBar.transform.localScale = HelthBar.transform.localScale - new
Vector3(0.0f, 0.0f, HP1*DamageHP);
    if (HP < 1)
    {
        Animator anim = gameObject.GetComponent<Animator>();
        anim.SetBool("Ded", true);
        Destroy(HelthBar);
        StartCoroutine(Die());
    }
}
```

Рисунок 3. Скрипт, обрабатывающий получение урона противниками

Для адаптации сцены к требованиям проекта требуется заменить тип камеры на XR камеру, что требует установки соответствующих библиотек, включая AR Core и AR Subsystems. Указанные библиотеки обеспечивают добавление на сцену XR камеры и расширенную возможность определения плоскостей в реальном мире. Отличительной особенностью XR камеры от стандартной 3D камеры является то, что первая не фиксирует объекты сцены и отображает только реальный мир, на который накладываются объекты из сцены.

Для обеспечения надлежащего качества и стабильности работы приложения был проведен smoke test, позволяющий быстро оценить функциональность основных компонентов программного продукта.

В процессе тестирования были выявлены и устранены ошибки, которые могли привести к неправильной работе приложения.

После завершения тестирования программный продукт был готов к использованию в качестве бета-версии, однако некоторые недочеты могут проявиться только после длительного использования на разных устройствах.

### **Заключение.**

В результате проектирования разработано мобильное приложение AR «Битва с роботами», которое может быть интересно широкой целевой аудитории, но, в первую очередь, пользователям в возрасте от 12 до 18 лет.

Для достижения поставленных целей были выполнены следующие задачи:

1. Рассмотрены аналоги мобильных игр с использованием дополненной реальности, проведен их сравнительный анализ, выявлены их достоинства и ограничения;
2. Проведено обоснование выбора среды разработки и используемых технологий, а также выбор языка программирования;
3. Реализовано приложение с использованием элементов AR;
4. Проведено тестирование мобильного приложения.

Заключительный отчет о разработке автоматизированной системы демонстрирует успешное прохождение тестов и достижение основных целей проекта. Было написано более 15 скриптов, реализующих различные функции приложения, с использованием официально поддерживаемых библиотек и методов.

Написание скриптов проводилось в интегрированной среде разработки Visual Studio, обеспечивая высокую степень структурированности и читаемости кода.

Приложение было успешно протестировано на различных устройствах, что говорит о его стабильной работе в наиболее возможных окружениях.

Однако, в дальнейшем рекомендуется проводить дополнительное тестирование приложения на различных устройствах с разной конфигурацией, чтобы обеспечить максимально возможную стабильность и надежность работы.

### **Список литературы**

- [1] <https://www.statista.com/outlook/amo/ar-vr/worldwide>
- [2] <https://newzoo.com/insights/articles/the-games-market-in-2021-the-year-in-numbers-esports-cloud-gaming>

## **AUGMENTED REALITY GAME «BATTLE WITH ROBOTS»**

***U.P. Kurilina***

*Student of the educational Institution "National Children's Technopark", student GUO "Gymnasium of Dobrusha"*

***D.A. Filon***

*Student of the Educational Institution "National Children's Technopark", a student of the State Educational Institution "Rechitsky District Lyceum"*

***M.S. Ilyasova***

*Master's student of the Department of Engineering Psychology and Ergonomics of BSUIR*

***Usenko F.V.***

*Master's student of the Department of Engineering Psychology and Ergonomics of BSUIR*

***L.R. Korkin***

*Master of Technical Sciences, Assistant of the Department of Engineering Psychology and Ergonomics of BSUIR*

***A.M. Prudnik***

*Associate Professor of Engineering Psychology and Ergonomics Department of BSUIR, Candidate of Technical Sciences, Associate Professor*

*Department of Engineering Psychology and Ergonomics  
Faculty of Computer-Aided Design  
Belarusian State University of Informatics and Radio Electronics, Republic of Belarus  
E-mail: aleksander.prudnik@bsuir.by*

**Abstract.** The components, objects, and structures necessary to create an AR application based on the Unity platform are considered. Some aspects of the development of the application on the real project "Battle with robots", written in the process of work, are also considered. In addition, the methods of testing the project during development and after the final assembly were applied in the work.

**Keywords:** VR technologies, augmented reality, Unity, C#, Git, Blender, Visual Studio.

УДК 336.74

## СОСТАВЛЕНИЕ РЕКОМЕНДАЦИЙ ДЛЯ МАГАЗИНОВ ТОРГОВОЙ СЕТИ НА ОСНОВЕ АНАЛИЗА ПОТРЕБИТЕЛЬСКОЙ КОРЗИНЫ



**К.А. Кот**

Магистрантка кафедры МиОЭС  
УО ГрГУ им. Янки Купалы  
kristinakot.gsl@gmail.com

### **К.А. Кот**

Окончила Гродненский государственный университет имени Янки Купалы. Магистрантка кафедры математического и информационного обеспечения экономических систем УО ГрГУ им. Я. Купалы.

**Аннотация.** Выполнен анализ магазина торговой сети и сделаны общие выводы на основе результатов предыдущих исследований магазинов из той же сети, составлены рекомендации для маркетинговой стратегии.

**Ключевые слова:** анализ потребительской корзины, ассоциативные правила, алгоритм Apriori.

### **Актуальность.**

В быстрорастущем мире торговли всегда являлся актуальным вопрос увеличения продаж и грамотного распределения ресурсов. Для этих целей были разработаны различные методы анализа потребительской корзины, такие как поиск последовательных шаблонов, кластеризация и сегментация, коллаборативная фильтрация, а также ассоциативные правила.

Ассоциативные правила являются важным методом анализа данных, особенно в области розничной торговли. Они используются для определения комбинаций товаров, которые часто покупаются вместе, а анализ потребительской корзины использует эти правила для оптимизации управления запасами, продвижения товаров и улучшения программ лояльности.

### **Применение ассоциативных правил.**

Основные направления для использования ассоциативных правил:

- **перекрестные продажи/продвижение:** анализируя данные, специалисты могут определить комбинации продуктов, которые часто покупаются вместе, и использовать эту информацию для совместного продвижения таких продуктов. Например, если клиент покупает обувь определенной марки, он также может быть заинтересован в покупке подходящих носков или шнурков для обуви. Совместно продвигая эти товары, магазины могут увеличить свои продажи;

- **управление запасами,** которое осуществляется путем выявления наиболее популярных продуктов/категорий, продаваемых в магазине. Если специалист замечает, что определенного товара часто нет в наличии, он может скорректировать свою стратегию закупок, чтобы быть уверенным, что у него достаточно запасов для удовлетворения покупательского спроса;

- **размещение товаров:** анализ потребительской корзины можно использовать для определения наиболее популярных категорий товаров в магазине и применять эту информацию для оптимизации размещения товара и оформления полок. Например, если анализ показывает, что определенный тип товара продается особенно хорошо, торговые точки могут выделить для

него больше места или расположить его на более видном месте. Это может помочь клиентам легче находить нужные им товары, что приведет к увеличению продаж.

### Анализ магазина с помощью алгоритма Apriori.

В качестве инструмента для выявления ассоциаций был использован один из алгоритмов поиска ассоциативных правил: Apriori. Объектом исследования является магазин торговой сети, исследования других магазинов из этой сети приведены в предыдущих исследованиях.

На основе данных описательной статистики были получены следующие результаты: в наборе данных присутствуют 531 продукт и 55332, максимальное количество товаров в чеке – 47, наиболее часто встречающиеся из них: хлеб; молоко; батон; сметана; колбаса.

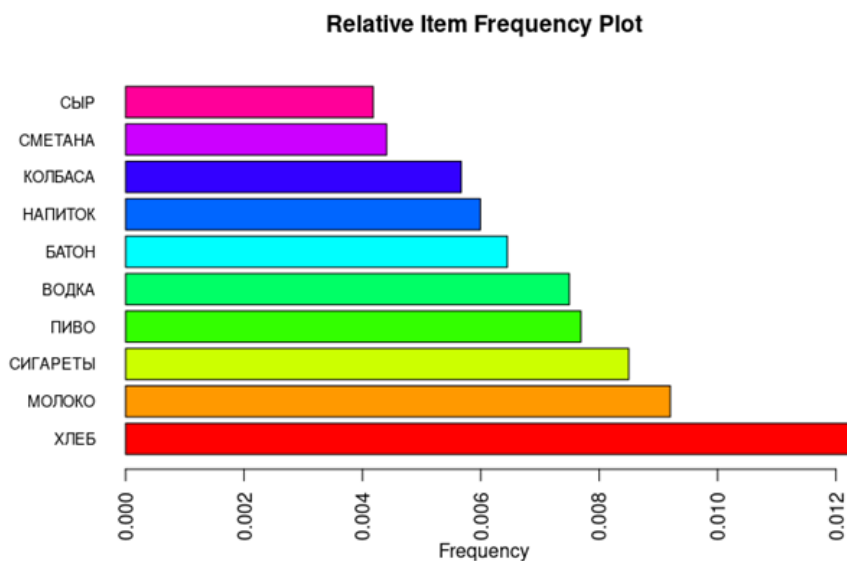


Рисунок 1. График наиболее часто покупаемых товаров

Применим алгоритм Apriori, передав в качестве входных параметров поддержку, равную 0,0001, и достоверность, равную 0,4.

lhs	rhs	support	confidence
[1]{БАТОН, МОЛОКО}	=> {ХЛЕБ}	0.001399938	0.5811518
[2]{БАТОН}	=> {ХЛЕБ}	0.003178239	0.4927650
[3]{СМЕТАНА}	=> {ХЛЕБ}	0.001959914	0.4445080
[4]{ТВОРОГ}	=> {МОЛОКО}	0.001407506	0.4439141
[5]{МАСЛО}	=> {ХЛЕБ}	0.001727852	0.4407979
[6]{БАТОН, ХЛЕБ}	=> {МОЛОКО}	0.001399938	0.4404762
[7]{МАЙОНЕЗ}	=> {ХЛЕБ}	0.001069503	0.4357657
[8]{КЕФИР}	=> {ХЛЕБ}	0.001470566	0.4331352
[9]{КЕФИР}	=> {МОЛОКО}	0.001455432	0.4286776
[10]{МОЛОКО}	=> {ХЛЕБ}	0.003937485	0.4279057
lhs	rhs	support	confidence
[11]{СЫР}	=> {ХЛЕБ}	0.001750554	0.4185766
[12]{ТВОРОГ}	=> {ХЛЕБ}	0.001306609	0.4120923
[13]{СМЕТАНА}	=> {МОЛОКО}	0.001806047	0.4096111

Как видно из полученных ассоциативных правил, ярко выделяется связь между молочными и хлебобулочными продуктами, по результатам lhs-rhs анализа (иными словами,



выявление того, что следует или предшествует покупке конкретного товара) наблюдается связь между покупкой майонеза и горошка/кукурузы, а также между сигаретами и соком:

lhs	rhs	support	confidence
[1] {ГОРОШЕК, СЫР}	=> {МАЙОНЕЗ}	0.58516405	0.5352113
[2] {БАТОН, ГОРОШЕК, ХЛЕБ}	=> {МАЙОНЕЗ}	0.54930614	0.5238095
[3] {КУКУРУЗА, ПАЛОЧКИ}	=> {МАЙОНЕЗ}	0.54706489	0.5121951
[4] {ГОРОШЕК, СМЕТАНА}	=> {МАЙОНЕЗ}	0.52051134	0.4909091
[5] {ГОРОШЕК, ЯЙЦО}	=> {МАЙОНЕЗ}	0.49154741	0.4181818
[6] {КУКУРУЗА, ГОРОШЕК}	=> {МАЙОНЕЗ}	0.46171700	0.4155844
[7] {БАТОН, СОК, ПЕЧЕНЬЕ}	=> {СИГАРЕТЫ}	0.5327534	0.8000000
[8] {ХЛЕБ, СОК, ПЕЧЕНЬЕ}	=> {СИГАРЕТЫ}	0.5149954	0.7631579
[9] {БАТОН, СОК, НЕКТАР}	=> {СИГАРЕТЫ}	0.5049948	0.6041667
[10] {СОК, ОГУРЦЫ}	=> {СИГАРЕТЫ}	0.4817174	0.6037736

Диаграмма рассеяния для 13 наиболее значимых правил представлена на рисунке 2:

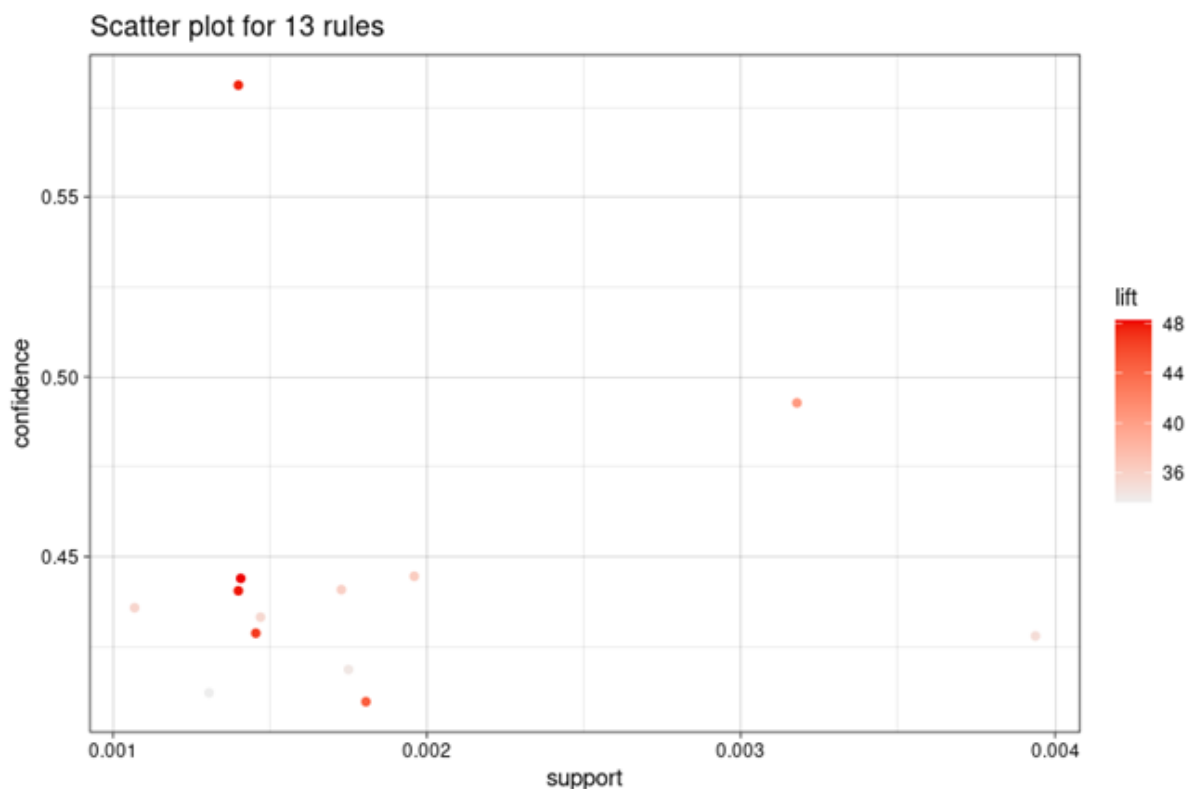


Рисунок 2. Диаграмма рассеяния

Для следующего графика (рисунок 3) были отобраны 9 правил с самым высоким значением достоверности. Более насыщенный цвет характерен для правил с самым высоким показателем подъемной силы:

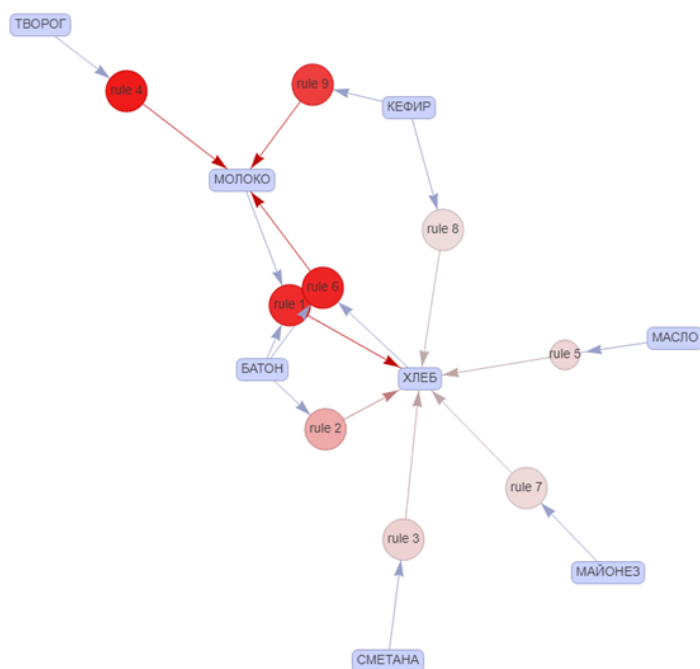


Рисунок 3. Представление правил в виде графа

На графике параллельных координат (рисунок 4) ярко выделяется связь между наличием в чеках молока, батона и хлеба:

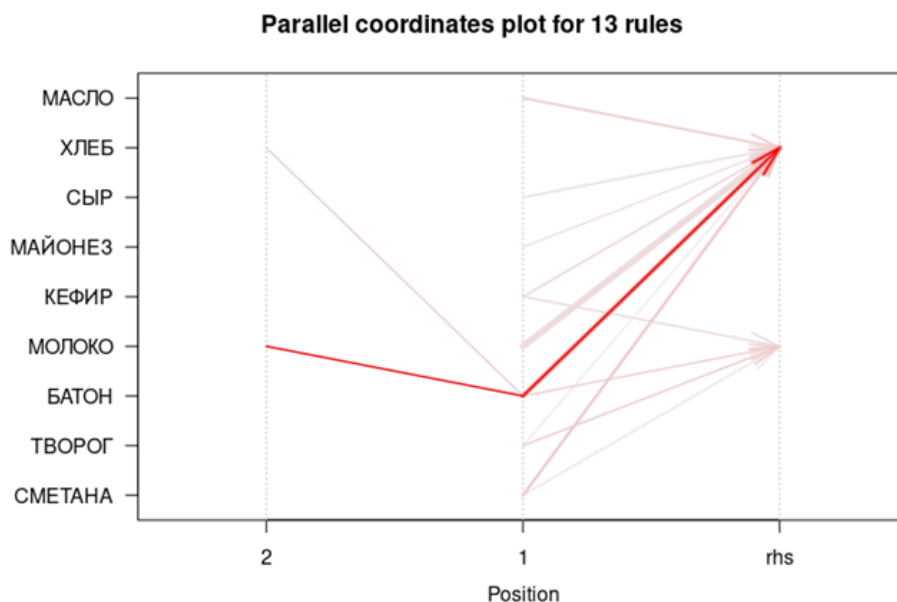


Рисунок 4. График параллельных координат

**Сравнение результатов с полученными ранее и составление рекомендаций.**

Исходя из полученных правил, прослеживается четкая тенденция совместных покупок молочной и хлебобулочной продукции, майонеза и консервированных круп, а также сигарет и сока. По результатам работы алгоритма в предыдущих исследованиях была также выявлена общая характерная черта - взаимосвязь между покупками молочной и хлебобулочной

продукцией. Кроме этого в них были выявлены правила, указывающие на наличие ассоциативной связи в покупках пива и чипсов, молочной продукции и сладостей, а также молочной и мясной продукции. Обобщив все полученные в ходе анализа результаты можно вынести несколько рекомендаций для данной торговой сети.

Поскольку во всех магазинах самыми часто попадающими в чеки товарами были хлеб и молоко, то за их наличием необходимо следить в первую очередь. Можно также располагать данные отделы рядом и ближе ко входу, ведь, как известно, если покупатель в течение 2-3 минут, после того, как зашел, приобретает один товар, повышается вероятность, что он купит что-либо еще. Располагать у входа можно, к примеру, только выпечку, а его молоко и его производные хранить подальше, поскольку из полученных ассоциативных правил мы знаем, что между их покупками наблюдается сильная ассоциативная связь. Кроме этого, продукцию из выявленных закономерностей можно использовать для совместного продвижения.

#### **Заключение.**

Во всех проанализированных магазинах выявлена ассоциативная связь между покупками молочной и хлебобулочной продукции, эти же товары являются наиболее часто приобретаемыми. Дополнительно можно отметить, что в некоторых магазинах сети наблюдались особенности, такие как: появление правил, указывающих на наличие взаимосвязи в покупках пива и чипсов, молочной продукции и сладостей, а также молочной и мясной продукции.

#### **Список литературы**

[1] Кот К. А. Алгоритм APRIORI и его принципы, лежащие в основе поиска ассоциативных правил / Наука-2022 [Электронный ресурс]: сб. науч. ст. / ГрГУ им. Янки Купалы; редкол.: А. Ф. Проневич (гл. ред.), А. С. Горный (зам. гл. ред.) [и др.]. – Объем электрон. дан. 12,4 Мбайт. – 62,0 уч.-изд. л. – Гродно: ГрГУ, 2022. – С. 688 – 693.

[2] Кот, К. А. Аффинитивный анализ данных потребительской корзины с помощью алгоритма Apriori / К. А. Кот, Н. В. Марковская // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник научных статей VIII Международной научно-практической конференции, Минск, 11-12 мая 2022 года / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. А. Богуш [и др.]. – Минск, 2022. – С. 234–239.

[3] Кот, К. А. Использование алгоритма Apriori для поиска ассоциативных правил в потребительской корзине / К. А. Кот; науч. рук. Н. В. Марковская // Студенческий вестник. – 2021. – № 47. Ч. 7. – С. 15-17.

[4] Кот, К. А. Использование алгоритма APRIORI для анализа потребительских корзин сети продовольственных магазинов / К. А. Кот; науч. рук. Н. В. Марковская // Студенческий вестник. – 2022. – № 47. Ч. 7. – С. 39-42.

## **MAKING RECOMMENDATIONS FOR STORES OF THE RETAIL CHAIN BASED ON THE MARKET BASKET ANALYSIS**

***K.A. Kot***

*Master student of the Department of  
Mathematical and Information Support of  
Economic Systems, YKSUG*

*Yanka Kupala State University of Grodno, Republic of Belarus*

*E-mail: kristinakot.gsl@gmail.com*

**Abstract.** An analysis of retail chains was carried out and general surveys were conducted based on the results of research of previous stores from the same retail chain, the recommendations for marketing strategy were made.

**Keywords:** market basket analysis, association rules, Apriori algorithm.

УДК 004.023

## ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ОПТИМИЗАЦИИ ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ В ЖИЛОМ ДОМЕ



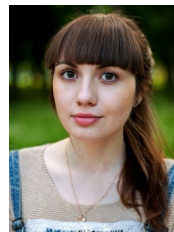
**Т.А. Васяева**

Декан факультета  
информационных систем и  
технологий, доцент кафедры  
автоматизированных  
систем управления ДОННТУ,  
кандидат технических наук,  
доцент  
vasyaeva@gmail.com



**С.В. Хмелевой**

Доцент кафедры  
автоматизированных  
систем управления  
ДОННТУ, кандидат  
технических наук,  
доцент  
hmelevoy.sergey@gmail.com



**А.С. Усова**

Магистр кафедры  
автоматизированных систем  
управления ДОННТУ  
anastasiasergeevna127@gmail.com

### **Т.А. Васяева**

Окончила Донецкий национальный технический университет. Область научных интересов связана с разработкой методов и алгоритмов интеллектуального анализа данных, организацией учебного и научно-исследовательского процессов в техническом университете.

### **С.В.Хмелевой**

Окончил Донецкий национальный технический университет. Область научных интересов связана с исследованием нейросетевых технологий, эволюционных вычислений, организацией учебного и научно-исследовательского процессов в техническом университете.

### **А.С. Усова**

Окончила Донецкий национальный технический университет. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем.

**Аннотация.** Рассмотрена проблема оптимизации потребления электроэнергии, проанализированы подходы к ее решению. Разработан генетический алгоритм применительно к Demand Response. Разработано техническое обеспечение интеллектуальной системы оптимизации потребления электроэнергии. Приведены результаты ее работы.

**Ключевые слова:** генетический алгоритм, фитнес-функция, Demand Response, оптимизация, интеллектуальные системы.

### **Введение.**

В последние годы количество устройств и технологий, которые потребляют энергию в жилых помещениях, значительно увеличилось. Это стало вызовом для производителей энергии и систем управления ею, поскольку потребление энергии в жилых помещениях может существенно влиять на окружающую среду и затраты на энергию. В условиях растущих цен на энергоресурсы и стремительного развития технологий, все больше людей задумываются о том, как можно сократить расходы на электроэнергию и одновременно не ущемить свой комфорт. Решение этой задачи может быть достигнуто благодаря оптимизации потребления электроэнергии в жилых помещениях с помощью искусственного интеллекта.

### **Актуальность.**

Энергопотребление является одним из основных источников выбросов углерода, что наносит вред окружающей среде и усиливает изменение климата. В связи с этим, снижение

энергопотребления становится все более важным в нашей жизни. Несмотря на то, что многие люди уже принимают меры для сокращения потребления энергии, все еще есть много возможностей для улучшения ситуации. Во многих странах для оптимизации потребления электроэнергии вводят программу Demand Response [1]. Разрабатываются системы оптимизации энергопотребления в жилых помещениях, которые могут помочь домовладельцам значительно сократить свой энергопотребление и, соответственно, уменьшить свой углеродный след.

#### **Анализ принципов работы систем оптимизации энергопотребления.**

Одним из последних инноваций в области оптимизации энергопотребления в жилых помещениях является смарт-технологии. Они позволяют собирать и анализировать данные о потреблении электроэнергии, газа и воды, чтобы оптимизировать использование ресурсов и снизить затраты на энергию.

Системы управления энергопотреблением также становятся все более интеллектуальными. Например, автоматические термостаты, оснащенные датчиками движения, могут определять, когда жильцы находятся дома, и регулировать температуру соответствующим образом.

Другим примером инновационных технологий являются интеллектуальные стиральные машины и посудомоечные машины. Они могут использовать информацию о потреблении энергии в режиме реального времени, чтобы оптимизировать свою работу и снизить затраты на энергию.

Продвинутые системы оптимизации энергопотребления также начинают постепенно устанавливаться в жилых помещениях, позволяя жильцам максимально использовать электроэнергию в те часы, когда цены на электроэнергию наиболее низкие.

Однако, несмотря на то, что эти технологии могут оказать значительное влияние на улучшение энергоэффективности в жилых помещениях, для их широкого внедрения необходимо совершенствование инфраструктуры.

#### **Постановка проблемы.**

Спрос на электроэнергию в жилых помещениях может значительно варьироваться в зависимости от времени суток, дня недели, сезона и других факторов. Оптимизация потребления энергии позволяет снизить спрос на электроэнергию в периоды пикового спроса, когда цены на электроэнергию наиболее высоки. Кроме того, оптимизация потребления энергии позволяет снизить нагрузку на энергетические системы, что в свою очередь уменьшает риск перегрузки системы и повышает ее надежность.

В работе решается задача минимизации расходов на электроэнергию и ограничение пиковой нагрузки потребления электроэнергии.

Наиболее значительным объектом оптимизации для жилого сектора являются домашние электроприборы. В рамках данной работы предлагается разделять электробытовые приборы на следующие категории: не оптимизирующийся, с термостатическим управлением, с циклическим режимом работы. Для них будут оптимизироваться временные моменты включения, для минимизации расходов на электроэнергию и снижение пиковой нагрузки на сеть. Формальная постановка задачи оптимизации описана нами ранее [2].

#### **Генетический алгоритм для решения задачи оптимизации потребления электроэнергии.**

Генетические алгоритмы [3] хорошо применимы для решения задач оптимизации. Многие авторы их использовали для оптимизации потребления электроэнергии [4]. Будем использовать ГА для разработки интеллектуальной системы в многоквартирном доме по управления электроприборами жильцов. В каждой квартире устанавливается микроконтроллер, который и будет управлять электроприборами. Пользователи дают информацию администратору о приборах, которые будут регулироваться. Наименование

электроприбора и его потребляемую мощность. Также пользователи составляют расписание (часы работы) для их электрооборудования в квартире. При этом предусмотрено, что у жильцов установлены трёхтарифные счетчики и оплата за электроэнергию начисляется в определённые часы по разным тарифам.

Пользователь в клиентском приложении может осуществлять ввод расписания для изменения, текущего состояние электроприборов в квартире. Или в программе производить осуществления команды на включение или выключение электроприбора непосредственно в тот же момент времени, рис. 1.

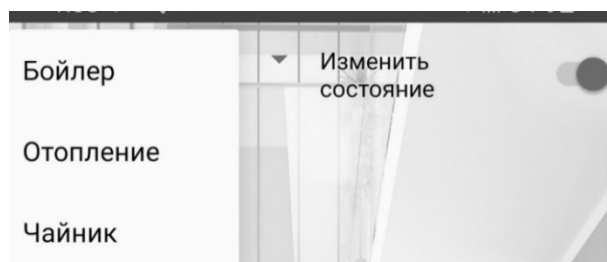


Рисунок 1. Фрагмент экранной формы «Включения электроприбора»

После все данные о командах с приложения передаются на Bluetooth модуль, этот модуль передаёт команды установленному микроконтроллеру в квартире, после чего они поступают на приемопередатчик физического уровня. И с помощью технологии Ethernet данные о командах поступят на сервер, а затем на ПК, где и будут обработаны. Входные данные представлены в табл.1.

Таблица 1. Входные данные для программного обеспечения

Обозначение	Описание
$M$	количество квартир в доме
$N$	количество управляемых электроприборов
$i$	индекс электроприбора
$t_i$	время включения $i$ -го электроприбора, где $i = [1, \dots, N]$
$\Delta t_i$	длительность времени работы электроприбора
$P_i$	мощность $i$ -го электроприбора, где $i = [1, \dots, N]$
$Vkl_{i,t}$	текущее состояние электроприбора, если $Vkl_i = 1$ – включен, если $Vkl_i = 0$ – выключен
$TimeCost_{i,t}$	текущая стоимость электроэнергии, потребленной в течение любого интервала времени $t$ , $t \in T$
$T$	горизонт планирования, $T = 24$
$EappUtil$	полезность бытовой техники (на сколько пользователю удобна работа прибора в конкретное время, измеряется от 0 до 1)
$EcostSavings$	экономия электроэнергии, выраженная в процентах

В процессе работы системы, основанной на выполнении ГА, будет выполняться пересчёт фитнес-функции и при необходимости отключатся электроприборы с наименьшим приоритетом, то есть произведутся ограничения по потребляемой мощности. На выходе системы оптимизации потребления электроэнергии, получаем оптимальное расписание работы электроприборов со снижением спроса, влекущее за собой

минимизацию суммарных затрат на электроэнергию, с учетом стоимости разных тарифов, а также снижением пиковой нагрузки потребления электроэнергии.

Формат выходных данных:

$$Y = \{ [N_i(\text{Электроприбор}), ] [VKL_{i,t}(\text{Текущее состояние электроприбора}), ] [t_i(\text{Часы работы}), ] [P_i(\text{Мощность}), ] [EcostSavings(\text{экономия электроэнергии})] \};$$

Генетический алгоритм будем использовать классический. Кодирование хромосомы в данной работе предлагается двоичное: хромосомы рассматриваются как массив битов, которые представляют состояние включения (равно 1) или выключения (равно 0) электроприборов. Интервал включения/выключения хромосом предлагается брать 10 минут, таким образом, состояние включения-выключения одного прибора в течение 24 часов можно закодировать  $24 \cdot 6 = 144$  битами. Дополнительно к этим битам в соответствии с работой [5] предлагается ввести дополнительные параметры, которые могут влиять на построение хромосомы. А именно, параметр определяющий id класса электроприбора; параметр для id электроприбора (порядковый номер электроприбора, закреплённый за ним в базе данных); параметр прерывания назначается для того чтобы определить может ли работа электроприбора прерываться или нет (1-да, 0-нет); параметр приоритета отводится для приоритетности электроприбора (1-да, 0-нет), табл. 2. Таким образом, общая длина хромосомы будет достигать 152 бит или 19 байт.

Таблица 2. Кодирование хромосомы

id класса	id электроприбора	Расписание	Прерывается работа или нет	Бит приоритета
3 бита	3 бита	144 бита - раз в 10 минут (6 бит на 1 час)	1 бит	1 бит

В соответствии с кодировкой хромосомы возможно использование стандартных генетических операторов – одноточечный кроссинговер, мутация, классическая репродукция. Ограничения в исследуемой работе, следующие: принудительное включение электроприборов пользователем; будет выставлен в 1 бит приоритетности и будет произведён перерасчёт ГА, исходя из текущего момента пересчитывается как дальше будет включаться электрооборудование; снижение пиковой нагрузки потребления; учет комфорта пользователя: своевременное периодическое включение приборов с циклическим режимом и поддержка работы приборов с термостатическим управлением в штатном режиме. Разработка фитнес-функции выполнена в [6]. Отметим что она состоит из двух составляющих: сэкономленная электроэнергия и полезность бытовой техники (на сколько пользователю удобна работа прибора в конкретное время, измеряется от 0 до 1).

#### Техническое обеспечение.

В разработанной системе оптимизации потребления электроэнергии различают клиентскую подсистему управления электроприборами, которая размещена в каждой квартире, и серверную подсистему централизованного расчета и оптимизации, которая размещена в серверной возле дома.

В клиентской подсистеме для каждой квартиры будет установлен:

- микроконтроллер Cortex M4 STM32F4Discovery, осуществляющий приём данных с Bluetooth модуля, обработку данных, выдачу управляющих сигналов;
- электрический многотарифный счётчик ВЕКТОР-3, предназначенный для учета пройденного через него количества электроэнергии, которая измеряется в Вт/ч;
- реле NRP-03K-C-05D-H, обеспечивающее вход 5 V и замыкание цепи 220 V;

–Bluetooth модуль HC-06, подключенный к микроконтроллеру, принимающий Bluetooth сигналы с мобильного устройства пользователя для дистанционного управления электроприборами, а затем осуществляющий передачу принятых данных о команде на микроконтроллер.

–приемопередатчик DP83848 PHYTER, подключенный к микроконтроллеру, при реализации Ethernet-интерфейса выступает трансивером физического уровня.

На рис. 2. представлена схема технического обеспечения клиентской подсистемы.

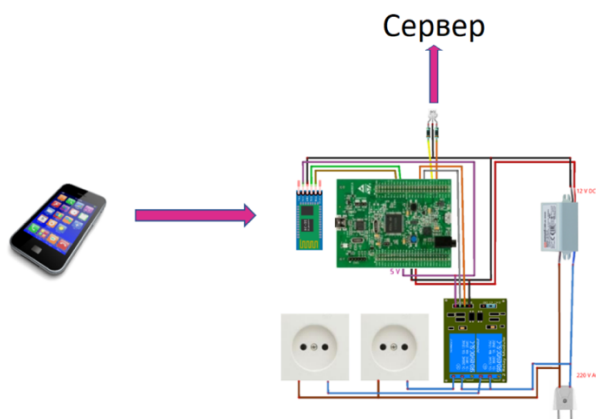


Рисунок 2. Схема технического обеспечения клиентской подсистемы

Для серверной подсистемы в серверной устан разрабатываемо ПО для оптимизации потребления электроэнергии, маршрутизатор TP-LINK Archer AX73 и коммутатор Huawei S1720X-32XWR.

Схема технического размещения оборудования изображена на рис. 3.

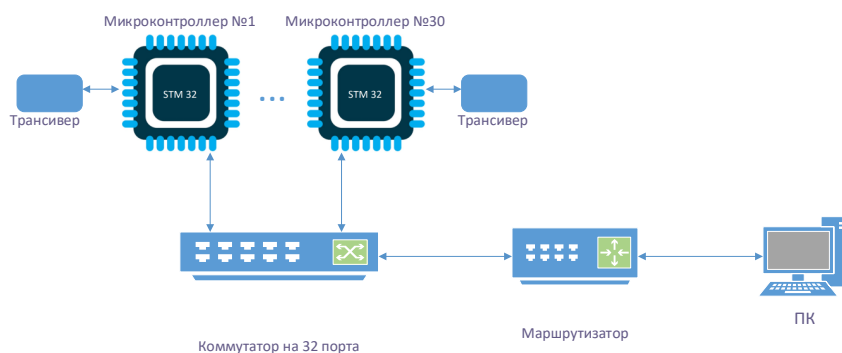


Рисунок 3. Схема технического размещения оборудования

### Результат работы

На рис. 4 изображен фрагмент неоптимизированного расписания, представляющее собой набор моментов включения и выключения электрооборудования. На рис. 5 изображён фрагмент работы ГА для оптимизированного расписания времени работы электроприборов в квартире. В табл. 3 представлен тестовый набор с описанием оптимального времени работы электроприборов.



```

id_кл id_эл расписание
3 0 1000111000 1110110110 1111111101 0101100111 1100100011 0000101101 1000100111 0001100011 0101110010 0000110000 1101000111 1111110010 0101101110 0111101111 1010 1 3
3 1 1001111000 1110100011 1001100000 0110101010 1010001001 1101111111 0100100101 1101111101 0111100011 0101011111 0010111001 1110011110 1010100111 1100110010 0111 1 3
3 2 0011111000 1110000100 1100010101 1001110000 1111101010 1101000001 1101010000 1110011100 1110000111 1100011110 1110010001 0100111111 0111101110 0100 1 3
3 3 0010101001 0000010111 0101101110 0111001011 0101111010 1101111111 1100101110 0100000011 0011100001 1010101111 0110001101 0111001000 1000101000 0100110010 1110 1 3
3 4 1111010100 1011000011 1011000100 0010110000 0101110101 0010000110 0010000110 0111100101 1010011111 0010111111 0111000000 0001100010 1111010101 0000011111 1010 1 3
3 5 1010000111 0110011001 1010110111 0001001011 0100100110 0100110011 0011110011 1011110101 0100111110 0100010111 1101101100 1110100011 0011010100 0010110011 0101 1 3
3 6 1000011110 1100100001 0001101001 0100110000 1011000100 1011111100 1001101101 1001101000 0011110100 0100000111 1101000001 1101000001 0000011111 0101010101 1100011110 1001 1 3
3 7 1011100001 1111111110 1101000010 0010100011 1001000100 0001000101 1011100101 1011011111 0110000110 1000010111 1101110011 1000110100 0001010110 1000010100 0001 1 3
2 8 1101101111 0011110001 1000010000 1100101100 0100001111 0111111100 0111010001 0010011000 1101110101 0011011011 1011001001 1101001001 0110011000 1000011110 0100 1 2
2 9 0001100010 1001001101 0010001101 0010110110 0111110111 1001000101 0111001010 0001111111 1010001000 0011001011 1011000111 0010100011 0010011001 1111001100 1000 1 2
2 10 1011011100 1100011100 0011000001 1011001100 1000101111 1110111000 1110111011 1111100001 0100100110 1001110111 0100011000 0010100010 0100001100 1100011111 1010 1 2
    
```

Рисунок 4. Фрагмент запуска программы работы ГА для построения графика работы электроприборов

```

id_кл id_эл расписание
3 0 1111111111 1111111111 1111111111 1101111111 1100000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 1111 1 3
3 1 1111111111 1111111111 1111111111 1101111111 1100000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000011 1111 1 3
3 2 1111111111 1111111111 1111111111 1100000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000011 1111 1 3
3 3 1111011111 1111111111 1111111111 1100000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000011 1111 1 3
3 4 1111111111 1111111111 1111111111 1100000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000000 0000000011 1111 1 3
3 5 0000000000 0000000000 0000000000 0000000000 0000000000 0000000001 1111111111 1111111111 1111111111 1100000000 0000000000 0000000000 0000000000 0000000000 0000 1 3
3 6 0000000000 0000000000 0000000000 0000000000 0000000000 0000000001 1111111111 1111111111 1111111111 1100000000 0000000000 0000000000 0000000000 0000000000 0000 1 3
3 7 1111111111 1111111111 1111111111 1100000000 0000000000 0000000000 0111111111 1111111111 1111111111 1100000000 0000000000 0000000000 0000000011 1111111111 1111 1 3
2 8 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001100011 0001 1 2
2 9 0000000000 0000000000 0000000000 0000011111 1100000000 0000000000 0000000000 0000000000 0000000000 0000011111 1100000000 0000000000 0000000000 0000000000 0000 1 2
2 10 0000000000 0000000000 0000000000 0000000000 1110101010 1010101010 0000000000 0000000000 0000000000 0000000000 0000000000 1110101010 1010101010 1010100000 0000000000 0001 1 2
    
```

Рисунок 5. Фрагмент программы работы ГА для построения графика работы электроприборов

Таблица 3. Результат работы ГА для построения графика работы электроприборов

id_кл	id_эл	Наименование	Время работы	Прерыва- ние	Приори- тет	
3-цикли- ческий	0	стиральная машина	только в ночной тариф	1	3	
	1	посудомоечная машина		1		
	2	умный пылесос (зарядка)		1		
	3	гироборд (зарядка)		1		
	4	сушилка		1		
	5	мультиварка		дневной тариф, но не в час пик		1
	6	хлебопечка		1		
2-термоста- тический	7	ноутбук (зарядка)	в любое время, кроме часа пик	1	2	
	8	холодильник	20 мин работает; 30 мин не работает	1		
	9	бойлер	вкл. на 70 мин перед часом пик	1		
	10	кондиционер	вкл. на 20 мин до часа пик; в часы пик 10 мин работает, 10 мин не работает	1		

Как видно из рис. 5 и табл. 3 система оптимизирует моменты включения и выключения электрооборудования, из-за чего происходит минимизация расходов на электроэнергию и снижение пиковой нагрузки потребления электроэнергии.

#### Заключение.

Интеллектуальное управление энергией – это эффективный способ снижения расходов на электроэнергию в жилых помещениях. В работе представлена интеллектуальная система оптимизации потребления электроэнергии в многоквартирном доме. Для оптимизации использован генетический алгоритм. Разработано техническое обеспечение системы.

### Список литературы

- [1] Nan S, Zhou M, Li G. Optimal residential community demand response scheduling in smart grid. Appl Energy 2018; 210:1280-9..
- [2] Усова А.С., Хмелевой С.В. Компьютеризированная система оптимизации потребления электроэнергии жилого дома // Информатика и кибернетика Донецк, ДонНТУ – 2020 – № 1(19) – С.27-35.
- [3] Javaid, N., Javaid, S., Abdul,W., Ahmed, I.; Almogren, A., Alamri, A., Niaz, I.A. A Hybrid Genetic Wind Driven Heuristic Optimization Algorithm for Demand Side Management in Smart Grid // Energies, Vol.10, N 319, 2017, - p.1-7.
- [4] Oladeji, O., Olakanmi, O. O. A genetic algorithm approach to energy consumption scheduling under demand response. In 2014 IEEE 6th International Conference on Adaptive Science & Technology (ICAST) (pp. 1-6). IEEE. – 2014
- [5] Javaid, N., Javaid, S., Abdul,W., Ahmed, I.; Almogren, A., Alamri, A., Niaz, I.A. A Hybrid Genetic Wind Driven Heuristic Optimization Algorithm for Demand Side Management in Smart Grid // Energies, Vol.10, N 319, 2017, - p.1-7..
- [6] Хмелевой, С. В. Генетический алгоритм для решения задачи оптимизации потребления электроэнергии в жилом доме / С. В. Хмелевой, А. С. Усова, О. В. Ченгарь // Информатика, управляющие системы, математическое и компьютерное моделирование (ИУСМКМ-2021): Материалы XII Международной научно-технической конференции в рамках VII Международного Научного форума Донецкой Народной Республики к 100-летию ДонНТУ, Донецк, 26–27 мая 2021 года. – Донецк: Донецкий национальный технический университет, 2021. – С. 35-40.

## INTELLIGENT SYSTEM FOR OPTIMIZING ELECTRICITY CONSUMPTION IN A RESIDENTIAL HOUSE

***T.A. Vasiaeva***

*Dean of the Faculty of Information  
Systems and Technologies,  
Associate Professor of the  
Department of Automated Control  
Systems of DONNTU, PhD of  
Technical Sciences, Associate  
Professor*

***S.V. Hmelevoy***

*Associate Professor of the  
Department of Automated Control  
Systems of DONNTU, PhD of  
Technical Sciences, Associate  
Professor*

***A.S. Usova***

*Master of the Department of  
Automated Control Systems,  
DONNTU*

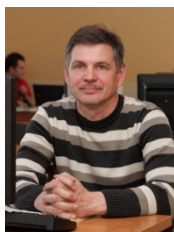
*Department of Automated Control Systems of DONNTU  
Faculty of Information Systems and Technologies  
Donetsk National Technical University, Russian Federation  
E-mail: vasyaeva@gmail.com*

**Abstract.** The problem of optimization of electricity consumption is considered, approaches to its solution are analyzed. A genetic algorithm has been developed for Demand Response. The technical support for an intelligent system for optimizing electricity consumption has been developed. The results of her work are presented.

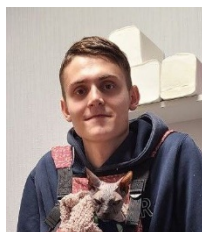
**Keywords:** genetic algorithm, fitness function, Demand Response, optimization, intelligent systems.

УДК 528.021.6/.7+681.2

## ПРИНЦИПЫ РАБОТЫ И СРАВНЕНИЕ ХАРАКТЕРИСТИК ТОЧНОСТИ СПУТНИКОВЫХ ПРИЕМНИКОВ



**М.М. Татур**  
Профессор кафедры  
электронных  
вычислительных машин  
БГУИР, доктор технических  
наук  
tatur@bsuir.by



**А.Д. Конигов**  
Аспирант кафедры  
электронных  
вычислительных машин  
БГУИР  
skonikov@mail.com



**И.Н. Носырев**  
Магистрант гр. 155841,  
кафедра ЭВМ, БГУИР  
nosureviluha@mail.ru

### **М.М. Татур**

Окончил Минское высшее инженерное зенитное ракетное училище ПВО в 1981. Область научных интересов связана с мобильной робототехникой и интеллектуальным анализом данных.

### **А.Д. Конигов**

Является аспирантом кафедры электронных вычислительных систем Белорусского государственного университета информатики и радиоэлектроники

### **И.Н. Носырев**

Является магистрантом Белорусского государственного университета информатики и радиоэлектроники по специальности 1-40 80 01 «Компьютерная инженерия (Хранение и обработка данных)».

**Аннотация.** Основой целью данной статьи является рассмотрение основных положений, лежащих в основе спутникового позиционирования. Приведены основные типы ошибок, влияющие на точность определения псевдодальностей, методы определения координат, типы спутниковых приемников. Описан текстовый протокол выдачи спутниковой информации NMEA-0813. Рассчитаны значения показателей точности позиционирования для двух моделей приемников.

**Ключевые слова:** GPS, GLONASS, RTK, NMEA, спутниковый приемник.

### **Введение.**

В настоящее время существуют и широко используются системы глобального позиционирования GPS, GLONASS, GALILEO, BEIDOU, QZSS, NavIC с различной государственной и ведомственной принадлежностью. Каждый пользователь мобильного телефона или навигатора используют функции системы позиционирования для решения своих утилитарных задач, не задумываясь о сложной технологической цепочке вычислений, предшествующей получению финального решения (координат). В настоящей работе мы рассмотрим основные характеристики навигационных приемников в контексте технологии глобального позиционирования.

Итак, независимо от вида системы, ее состав включает: космический сегмент; сегмент управления; сегмент пользователей.

На примере GPS, космический сегмент системы, по состоянию на 07 апреля 2020 года использует 31 спутник и 1 выведен на техобслуживание. Спутниковая группировка обращается вокруг Земли по круговым орбитам с постоянным периодом обращения (11 часов 58 минут) для всех спутников. Таким образом, спутник совершает два витка вокруг Земли за одни звёздные сутки (23 часа 56 минут). Спутники излучают специальные радиосигналы.

Сегмент управления системы GPS состоит из станций слежения, расположенных по всему миру, которые обрабатывают данные от спутников и соотносят эти данные с математическими моделями орбит для каждого из спутников. В соответствии с моделями рассчитываются точные параметры орбит – эфемериды и коррекции часов для каждого спутника. Эта информация посылается на спутник, и через него на приемники GPS. Навигационное сообщение передается с частотой 50 Hz, и содержит параметры орбиты спутников GPS, коррекции времени, параметры модели ионосферы и прочие данные. Данные передаются на несущих L1 и L2.

Сегмент пользователей - это конечные потребители, использующие спутниковые приемники для позиционирования. Приемники получают сигналы со спутников для расчета своего местоположения, времени и скорости перемещения. GPS-приемники используются для позиционирования, коррекции времени и других навигационных целей.

### **1. Принципы измерения расстояний от спутников до приемника на примере системы GPS.**

Вычислению координат приемника предшествуют вычисления расстояний от спутника до приемника. Существуют две методики измерений расстояний: *кодовая* и *фазовая*.

#### **Кодовые измерения:**

Каждый спутник GPS транслирует свою уникальный псевдослучайную последовательность нулей и единиц (Pseudo Random Noise - PRN). Код C/A повторяется через каждые 1023 бита раз миллисекунду, то есть транслируется с частотой 1.023 MHz.

Приемник генерирует такие же последовательности для каждого из спутников (на соответствующих каналах), но не синхронно. Далее, приемник “сдвигает” свой код пока тот не совпадет с кодом спутника. Данный сдвиг и определяет время прохождения сигнала от спутника до приемника. Умножая время прохождения сигнала на скорость света получаем искомое расстояние. Так как существуют ряд погрешностей при определении расстояния по коду, это измеренное расстояние называется псевдодальностью.

#### **Фазовые измерения:**

По аналогии с кодовым сигналом на спутнике и приемнике генерируются фазомодулированные сигналы, только они никак не синхронизированы. Вычислением смещения фазомодулированного сигнала, переданного со спутника на приемник, определяется дробная часть фазы. Вычислением смещения фазомодулированного сигнала, переданного со спутника на приемник, определяется дробная часть фазы.

Основная проблема – найти целое число длин волн между спутником и приемником, так как волны не отличаются друг от друга. Задача по поиску числа целых фаз называется *решением фазовой неоднозначности*.

При определении дальности возможны различные источники измерительных погрешностей.

#### **Погрешность синхронизации времени.**

На спутниках установлены сверхточные цезиевые (атомные) часы. Однако даже такие сверхточные часы допускают каждые 3 часа погрешность в 1 миллиардную секунды. Остаточные погрешности спутниковых синхронизаторов достигают порядка нескольких наносекунд, что вызывает погрешность в определении расстояния около метра. Для устранения возникающих отклонений, спутниковые синхронизаторы постоянно проверяются наземными станциями и сверяются с генератором главных тактовых импульсов, который представляет из себя систему 10-ти сверхточных атомных часов. Погрешности и отклонения спутниковых синхронизаторов высчитываются и отправляются вместе с сообщениями, передаваемыми спутниками. При вычислении расстояния до спутников GPS-приемники вычитают погрешности спутниковых синхронизаторов из зафиксированного времени передачи, чтобы определить реальное время перемещения сигнала. Очевидно, что встраивать в приемники точные атомные часы – нереально. Однако если измерить расстояния до всех спутников одновременно, то для всех спутников будет одна и та же погрешность в определении расстояний. Таким образом,

можно принять погрешность синхронизатора приемника за неизвестную, которую необходимо найти. Для нахождения этой поправки необходимо принимать сигнал от четырех спутников, при этом можно использовать в приемниках обычный синхронизатор.

#### **Погрешность определения координат спутника.**

Точность в определении местоположения приемника зависит от того, насколько точно известно положение спутников. Орбиты спутников непрерывно контролируются с расположенных по всему свету наземных станций слежения. Их прогнозы орбит передаются на спутники, которые в свою очередь передают эту информацию на приемники. Точность предсказания орбитальной информации составляет несколько метров.

#### **Погрешность ввиду задержки в распространения сигнала в ионосфере.**

Скорость света зависит от состояния атмосферы. Верхние слои атмосферы (ионосфера) содержат заряженные частицы, которые замедляют скорость света и увеличивают частоту несущей. В некоторых приемниках используют математическую модель эффекта ионосферы для вычисления поправки. При получении приблизительной информации о концентрации заряженных частиц в ионосфере (передаваемой со спутников), эффект воздействия ионосферы можно снизить до 50%. Однако оставшаяся погрешность велика.

Воздействие ионосферы на электронные сигналы зависит от частоты сигнала. Чем выше частота, тем меньше воздействие. Поэтому, если передаются пакеты сигналов одновременно на двух частотах, то ионосфера может задержать код одной частоты, к примеру, на 4 метров, а на другой – на 5 метров. Измерить величину этих задержек не представляется возможным, однако можно измерить их разницу через различие времени прибытия сигналов, что в данном случае означает 1 метр расстояния между ними. Подставив ее в известную формулу зависимости ионосферной задержки от частоты сигнала, можно устранить эффект ионосферы. Поэтому, точные приемники для устранения эффекта ионосферы отслеживают оба сигнала с несущими L1 и L2. А простые – только сигнал L1. Это - основное отличие различных видов приемников.

#### **Погрешность ввиду задержки в распространения сигнала в тропосфере.**

Нижние слои атмосферы, содержащие водные пары, называются тропосферой. Они могут задерживать получения кода, и несущей. Эффект от тропосферы нельзя устранить, используя принцип двух частот. Единственный способ устранения погрешности из-за эффекта тропосферы – измерение содержания водяных паров, температуры и давления и использование специальных математических моделей для вычисления тропосферной задержки сигнала.

#### **Погрешность ввиду переотражения сигнала.**

При определении расстояния до спутника предполагается, что сигнал спутника поступает от спутника к антенне приемника по прямой. Однако наряду с прямым сигналом возможны и отраженные сигналы - от земли и предметов, расположенных рядом с антенной, которые также достигают антенны и создают помехи основному сигналу. Если путь переотраженного сигнала значительно длиннее прямого (более, чем на 10 метров), это позволяет разделить сигнал на пакеты и в значительной мере устранить через обработку сигнала.

## **2. Методы расчета координат приемника**

Существует несколько методов определения координат с помощью в глобальной системе позиционирования, среди которых выделяют автономные и относительные.

Суть автономного метода заключается в следующем:

- вычисляются кодовые псевдодалности до четырех и более спутников;
- на основе данных эфемерид рассчитывается координаты каждого из этих спутников. решается следующая система уравнений (приведена для четырех спутников):

$$\begin{aligned}(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 &= (ps_1 + r)^2 \\(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 &= (ps_2 + r)^2 \\(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2 &= (ps_3 + r)^2 \\(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2 &= (ps_4 + r)^2\end{aligned}$$

где  $(x_i, y_i, z_i)$  – координаты  $i$ -ого спутника,  $ps_i$  - кодовая псевдодальность  $i$ -ого спутника. Решая систему уравнений, получим значения искомым неизвестными  $(x, y, z, r)$  - три координаты и отклонение часов приемника.

Точность определения координат в автономном режиме невысока и может достигать десятков метров.

Относительные методы предполагают наличие других приемников (или приемника), антенны которых установлены в точке с заранее известными точными координатами. Такие приемники называются базовыми. Приемник координаты антенны которого необходимо вычислить называется ровером.

Предполагается, что погрешности спутникового синхронизатора, эфемерид, ионосферы, тропосферы с одинаковой силой влияют как на базовый приемник, так и на ровер из-за их близкого расположения.

Относительные методы можно разделить на:

- методы, использующие кодовые измерения
- методы, использующие фазовые измерения

Относительные методы получения координат в англоязычной литературе называются методами DGPS (differential global positioning system).

Методы DGPS можно разделить на:

- методы, работающие в режиме постобработки
- методы, работающие в режиме реального времени

Режим постобработки предусматривает совместную обработку данных от базового приемника и ровера на ЭВМ после проведения измерений.

Методы же реального времени предполагают передачу поправок через радиоканал от базовой станции на ровер. Поправки обычно передаются с периодом в несколько секунд

Относительный метод реального времени сантиметровой точности является метод RTK (real time kinematic).

### **3. Аппаратура пользователей систем глобального позиционирования.**

Типовой «приемник GPS» включает: антенну, приемное устройство, управляющее устройство (контроллер) и аккумулятор.

Антенна может быть встроена в корпус самого приемника или использоваться в виде выносного модуля, который соединен с приемником специальным кабелем. Антенны могут быть одночастотные, то есть настроенные на прием только частоты L1 или двухчастотные, принимающие частоты L1 и L2.

Все приемники являются многоканальными, где каждый канал “следит” за своим спутником. Чем больше каналов, тем легче приемнику обнаружить необходимое количество спутников.

Приемники можно классифицировать на:

- односистемные, ориентированные на прием сигналов только от одной навигационной системы (обычно GPS)
- многосистемные, принимающие сигналы одновременно от нескольких системы (GPS, GLONASS и т.д.).

По видам принимаемых сигналов приемники можно разделить на (для системы GPS):

- кодовые, одночастотные работающие по SA-коду
- кодово-фазовые одночастотные, обрабатывающие SA-код и фазовые измерения на частоте L1
- кодово-фазовые двухчастотные, использующие по коду и фазе на обоих частотах (L1, L2).

Приемники могут выдавать следующую информацию, получаемую в результате обработки спутниковых сигналов.

- координаты местоположения антенны приемника;

- скорость перемещения приемника (антенны приемника) и его азимут;
- точное время;
- информацию о спутниках;
- значение DOP (геометрическое ухудшение точности) для множества спутников, участвующих в решении.

Протоколы, используемые приемниками, обычно разделяют на бинарные и текстовые. Одним из самых распространённых текстовых протоколов является протокол NMEA.

#### 4. Протокол NMEA.

Для обмена данными между приёмниками спутниковых систем GPS, GLONASS, GALILEO, BEIDOU, QZSS, NavIC с другими модулями систем навигации и управления мобильных роботов используют стандарт NMEA 0183.

NMEA (National Marine Electronics Association) — это американская торговая организация, занимающаяся вопросами стандартизации протоколов коммуникации между оборудованием морской электроники. NMEA предлагает 4 основных стандарта:

- NMEA 0400 — стандарт, описывающий требования к монтажу и подключению электрооборудования.

- NMEA 0183 – комбинированный стандарт, описывающий как электрические требования, так и коммуникационные (формат данных) для взаимодействия с блоками морской электроники такими, как эхолоты, сонары, анемометры, гирокомпасы, системы автопилотирования, приёмниками спутниковых систем и многими другими на скорости 4800бод/с либо 38.4кбод/с для версии High Speed. Основан на последовательной шине UART/USART RS-422 и RS232.

- NMEA 2000 – коммуникационный стандарт, основанный на шине CAN, высокоскоростной, до 250кбит/с, используется для подключения дисплеев, систем телеметрии судовых двигателей, автопилоты, устройств измерения скорости ветра, акустических измерителей глубины и многих других.

- NMEA OneNet – самый высокоскоростной стандарт коммуникации, базирующийся физически на стандарте IEEE 802.3 Ethernet и протоколно на IPv6, предназначен для подключения радаров, передачи видео- и аудиоинформации.

Наиболее широко распространённым за пределами судового электрооборудования стал NMEA 0183, пришедший на смену NMEA 0180 и 0182. Последняя версия стандарта на текущий момент 4.11, вышедшая в ноябре 2018 года. Его популярность обусловлена широким применением в качестве основного протокола обмена данными с различными модулями спутниковой навигации.

Электрически стандарт базируется на RS-422 и является совместимым с RS-232. На уровне данных обеспечивается скорость 4800 бод/с, 8 бит данных, 1 стоп бит, без битов чётности и рукопожатий (для модификации NMEA-0183HS доступна скорость 38400бод/с). Используется типовой набор ASCII символов. Протокол однонаправленный, поэтому допускается к одному источнику, например, GPS-приёмнику, подключать несколько приёмников (слушателей), например, систему автопилота, систему отображения координат на дисплей и т.д.

Все передаваемые данные являются ASCII символами. Зарезервированными являются следующие символы:

<CR> – (0xd) возврат каретки;

<LF> – (0x0a) перенос строки, конец сообщения;

! – (0x21) разделитель начала обёрнутого сообщения;

\$ – (0x24) стартовый разделитель;

\* – (0x2a) разделитель контрольной суммы;

, – (0x2c) разделитель полей;

\ – (0x5c) разделитель полей тэгов;

^ – (0x5e) кодовый разделитель для нестандартных (локальных языковых символов);

~ – (0x7e) зарезервировано.

Максимальная длина сообщения составляет 82 символа, включая стартовый «!» или «\$» и заканчивая <LF>.

Пакет начинается символом «\$» или «!» (для пакетов с нестандартными символами).

Затем идут 5 символов, описывающих передатчик сообщения, из которых первых 2 символа — идентификатор передатчика, затем 3 символа — тип пакета. Далее идут поля данных, разделённые запятыми, если какое-то поле отсутствует, то после него ставится также запятая, т.е. оставляется пустым «,».

После полей данных ставится символ «\*» и за ним следует контрольная сумма в виде двухзначного шестнадцатеричного числа, рассчитываемого из всех предыдущих символов посредством функции исключающего или (XOR). Завершается сообщение парой символов возврата каретки и переноса строки <CR> и <LF>.

Рассмотрим следующий пример:

```
$GPGGA,195014.000,5391.788,N,02759.37,  
E,4,13,1.00,295.144,M,29.200,M,0.10,0000*40
```

Где GP – идентификатор устройства, в данном случае GPS приёмник, для ГЛОНАСС – GL, Galileo – GA, BeiDou – BD или GB, GPS + ГЛОНАСС – GN [6].

GGA – тип данных, в этом примере фиксированные данные глобальной системы позиционирования, возможны ещё:

GLL – географическое положение широта и долгота;

GSA – GNSS DOP и подробные сведения об активных спутниках;

GSV – GNSS спутники и подробная информация такая, как азимут и угол над горизонтом каждого спутника;

RMC – рекомендуемые минимально определяемые данные GPS;

VTG – курс относительно земли и скорость;

GST – приблизительная точность по вертикали и горизонтали.

195014.000 – время в формате UTC чммсс.ссс, где ч — часы, м — минуты, а с — секунды с миллисекундами.

5391.788 – широта, 2 цифры градусы, 2 цифры целая часть минут и дробная после точки переменной длины, в зависимости от точности. N – для северной широты и S ставят для южной.

02759.37 — долгота (3 цифры для градусов, 2 — для целой части минут и дробная часть переменной длины после точки. E – для восточной, W – для западной.

4 — идентификатор качества данных, 1 — данные некорректны, 2 — дифференциально скорректированная координата (для систем DGPS, WAAS и т. д.), 4 — RTK с фиксированной координатой (сантиметровая точность), 5 — RTK движущаяся.

13 — число спутников, используемых для определения координаты.

1.00 — параметр снижения точности (HDOP - horizontal dilution of precision).

295.144 – высота антенны приёмника над уровнем моря.

M – определяет систему счисления высоты M — метры, F – футы.

29.200 — геоидальное разделение, с его помощью можно получить высоту над эллипсоидом (NAE), для этого этот параметр надо вычесть из высоты.

M — система счисления геоидального разделения (метрическая).

0.10 — возраст коррекции (сколько времени прошло с последней коррекции данных).

0000 — идентификатор станции коррекции.

\*40 — разделитель «\*» и значение контрольной суммы в шестнадцатеричном формате.

В примере рассмотрен лишь один из возможных видов пакета, в то время, как их существует несколько, содержимое которых определяется идентификатором пакета и типом устройства (первые 5 символов после стартового \$ или !), например скорость движения, данные компаса, картографические данные, контрольные данные приёма маяков, маршруты и т.д. Более того, у многих производителей навигационного оборудования есть собственные, проприетарные типы пакетов, в которые может закладываться различная сервисная и служебная информация.



Стандартизация обмена данными позволяет добиться необходимых параметров надёжности и безопасности изделий, а также снизить временные и финансовые затраты на разработку систем, поскольку позволяет тиражировать собственные программные и аппаратные наработки, а также обеспечивает совместимость систем различных производителей.

### 5. Оценка точности определения координат GPS приемником

Основные меры, используемые для оценки точности определения координат антенны приемника - стандартное отклонение -  $SD$  (standard deviation) и круговое вероятное отклонение -  $CEP$  (circular error probable).

Стандартное отклонение определяется как квадратный корень из дисперсии случайной величины.

На практике распределение случайной величины неизвестно и имеется лишь конечная выборка значений случайной величины (в данном случае координаты точки). Тогда можно применить следующую формулу для расчета стандартного отклонения на основе выборки из  $N$  измерений:

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

где  $x_i$  -  $i$ -ое значение случайной величины,  $\bar{x}$  - среднее значение случайной величины.

Величина среднеквадратического отклонения показывает разброс значений в представленном множестве относительно средней величины. Таким образом, малое значение  $SD$  свидетельствует о том, что значения в множестве  $N$  компактно сгруппированы вокруг среднего значения. Круговое вероятностное отклонение определяется как значение радиуса круга, в который “попадает” заданный процент точек. Например, процент может быть задан, как 50%, 90%, 95%, 98%. Поставим эксперимент по оценке точности определения координат двух различных приемников: SIM808 и Ublox NEO-7M. Пусть оба приемника работают в автономном режиме определения координат. Антенны для обоих приемников установлены на расстоянии 15 сантиметров друг от друга. Оба приемника выдавали сообщения GPGLL в формате NMEA с частотой 1 Герц на PC с OS Windows 10. Для обработки сообщений от приемника Ublox NEO-7M на PC установлена программа U-Center 8.29, которая принимала сообщения и записывала их в файл. Аналогично, приемник SIM808 посылал сообщения на тот же компьютер, а программа Termite-3.3 принимала сообщения и записывала их в файл. Собранные файлы были сконvertированы из формата nmea (текстовый файл с nmea сообщениями) в формат grx с помощью онлайн сервиса [NMEA to GPX Converter Online - MyGeodata Cloud](https://www.mygeodata.cloud/nmea-to-gpx/).

Оба файла в формате grx были загружены в программу U-Center 8.29 для получения карт разброса, приведенных ниже на рисунках 1 и 2.

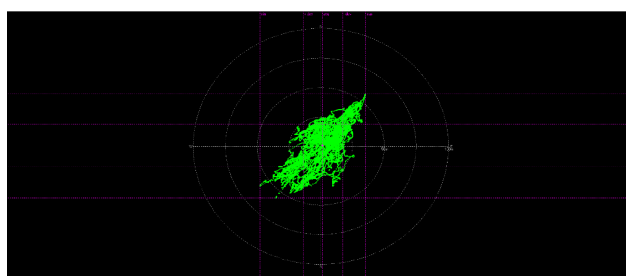


Рисунок 1. Карта измеренных координат, построенная в программе U-Center 8.29 для приемника SIM808

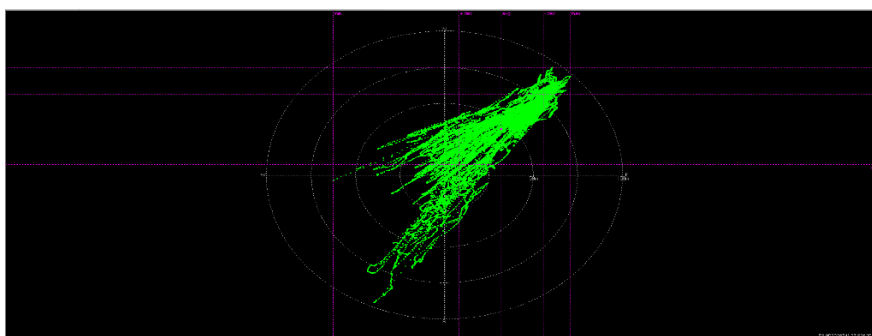


Рисунок 2. Карта измеренных координат, построенная в программе U-Center 8.29 для приемника U-Blox NEO7M.

По рисункам визуально можно оценить, что координаты, измеренные приемником SIM808 имеют меньший разброс. Для количественной оценки точности измерения координат можно воспользоваться программой DNR GPS. Результаты вычислений сведены в нижеприведенной таблице.

Таблица 1. Значения показателей точностей для приемников SIM808 и U-Blox NEO-7M

	SIM808	U-Blox NEO-7M
SD (широта)	0,000161	0,000218
SD (долгота)	0,000242	0,000363
СЕР (50%)	18,9	26,2
СЕР (90%)	37,7	47,5
СЕР(95%)	41,1	62,1
СЕР(98%)	44,7	83,5

Как следует из таблицы, все значения SD и СЕР приемника SIM808 меньше чем соответствующие показатели для приемника U-Blox NEO-7M, следовательно, можно сделать вывод, что приемник SIM808 измеряет координаты точнее, чем приемник U-Blox NEO-7M. Обладая такими данными, при необходимости, не составит труда определить на сколько эффективнее один приемник по сравнению с другим по каждому из показателей. По аналогии можно проводить исследования по оценке точности определения координат относительными методами, скорости перемещения, азимута и др. параметров, определяемых различными приемниками.

### Список литературы

- [1] Серапинас Б.Б. “Глобальные системы позиционирования” Москва, 2002.  
 [2] Антонович К. М. “Использование спутниковых радионавигационных систем в геодезии” Том 1 Москва, 2005.  
 [3] Яценков В.С. “Основы спутниковой навигации. Системы GPS NAVSTAR и ГЛОНАСС” Москва, 2005  
 [4] NMEA-0183 messages: Overview [Электронный ресурс]. URL: [https://receiverhelp.trimble.com/alloy-gnss/en-us/NMEA-0183messages\\_MessageOverview.html](https://receiverhelp.trimble.com/alloy-gnss/en-us/NMEA-0183messages_MessageOverview.html) (Дата обращения: 19.03.2023).  
 [5] Special Topics - GPS [Электронный ресурс]. URL: [https://www.youtube.com/watch?v=16xHIBmul\\_o&list=PLX2gX-ftPVXXGdn\\_8m2HCIJS7CfKMCwol](https://www.youtube.com/watch?v=16xHIBmul_o&list=PLX2gX-ftPVXXGdn_8m2HCIJS7CfKMCwol) (Дата обращения: 20.03.2023).

## **PRINCIPLES OF OPERATION AND COMPARISON OF ACCURACY CHARACTERISTICS OF SATELLITE RECEIVERS**

***M.M. Tatur***

*Professor of Electronic Computing  
Machines Department Professor,  
Doctor of Technical Science*

***A.D. Konikov***

*Postgraduate of Electronic  
Computing Machines Department*

***I.N. Nosarau***

*Master's student of group 155841,  
Electronic Computing Machines  
Department, BSUIR*

*Department of Electronic Computing Machines*

*Faculty of Computer Systems and Networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

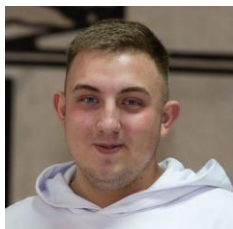
*E-mail: tatur@bsuir.by*

**Abstract.** The main purpose of this article is to consider the main provisions underlying satellite positioning. The main types of errors affecting the accuracy of determining pseudo-distances, methods of determining coordinates, types of satellite receivers are given. The text protocol for the issuance of satellite information NMEA-0813 is described. The values of positioning accuracy indicators for two receiver models are calculated.

**Keywords:** GPS, GLONASS, RTK, NMEA, satellite receiver.

УДК 004.774:656.131.7

## ПРОГРАММНАЯ ПОДДЕРЖКА УПРАВЛЕНИЯ ВЗАИМОДЕЙСТВИЕМ С КЛИЕНТАМИ В СФЕРЕ АРЕНДЫ АВТОМОБИЛЕЙ



**В.А. Миленцев**  
Магистрант кафедры  
ИПиЭ, ассистент  
кафедры  
экономической  
информатики БГУИР  
milvova864@gmail.com



**А.М. Прудник**  
Доцент кафедры ИПиЭ,  
кандидат технических  
наук, доцент  
aleksander.prudnik@bsuir.by



**И.В. Андриалович**  
заместитель декана  
факультета  
компьютерного  
проектирования БГУИР,  
соискатель  
andryinna@bsuir.by



**А.В. Воробей**  
аспирант кафедры  
ИПиЭ,  
ассистент  
кафедры ИПиЭ,  
магистр  
технических наук  
vorobey@bsuir.by

### **В.А. Миленцев**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с изучением интерфейсов информационных систем, пользовательских интерфейсов, front-end web development, UX-дизайном, а также с интеграционным тестированием.

### **А.М. Прудник**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с изучением взаимодействия человека с компьютером, интерфейсов информационных систем, пользовательских интерфейсов, front-end web development, UX-дизайном, а также с моделированием прохождения электромагнитных и акустических волн через вещество.

### **И.В. Андриалович**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием проблем психологического выгорания профессорско-педагогического состава учреждений высшего образования.

### **А.В. Воробей**

Окончила Белорусский государственный педагогический университет имени М. Танка» по специальности Биология и химия. В 2020 году окончила магистратуру БГУИР по специальности «Психология труда, инженерная психология, эргономика», выполняет исследования в рамках магистерской диссертации на тему «Система обеспечения безопасности труда при воздействии инфракрасного излучения на физиологические параметры человека». Обучается в аспирантуре БГУИР по специальности «Психология труда, инженерная психология, эргономика».

**Аннотация.** Описывается веб-приложение для фирмы по прокату автомобилей. Данное программное средство является источником актуальной и полезной информации для потенциальных арендаторов, а также платформой для размещения объявлений менеджерами по аренде автомобилей.

**Ключевые слова:** аренда транспортных средств, менеджер, клиент, веб-приложение, Java, MySQL.

### **Введение.**

Современная жизнь очень динамична, ее ритм заставляет человека двигаться с ней в ногу и максимально соответствовать ее быстро изменяющимся условиям.

Мобильность, энергичность, пунктуальность из обычных положительных качеств человека превратились в необходимые требования, которым он должен соответствовать, чтобы успешно существовать в современном мире.

Технические средства и удобные услуги современного рынка: «аренда авто» или «прокат машин» – эти незаменимые помощники человека, без которых уже немислимо само существование в нашем сложном и динамичном мире, только они помогают эффективно решать вопросы мобильности, динамичности и высокой производительности.

Привыкнув к этим неотъемлемым элементам современной жизни совсем недавно, в особенности к услуге – «аренда автомобилей» или, как часто пишут в объявлениях — «аренда и прокат автомобилей», человек уже не мыслит самого существования без них и испытывает чувство огромного дискомфорта, когда по каким-либо причинам утрачивает возможность пользоваться хоть одним из этих важных элементов, а в особенности – автомобилем, на что современный рынок услуг, предлагает одну из самых удобных из них — аренда машин, а точнее – прокат автомобилей без водителя.

Но не всегда покупатель останавливает свой выбор на одной компании-арендодателе. Зачастую он выбирает из нескольких вариантов.

В данном случае не совсем удобно следить за несколькими ресурсами, которые в то же время имеют разную структуру, наполненность и информативность для потенциального арендатора автомобиля, а также это делает некомфортным сравнение цен на аренду у разных компаний и поиск оптимального по цене предложения. Многие клиенты отдадут предпочтение сайтам, предоставляющим возможность связаться с менеджером напрямую и выяснить все детали в режиме онлайн.

Из-за недостатка ресурсов с таким набором функций арендаторам приходится совершать звонки либо писать письма по электронной почте, что влечёт за собой немалые потери времени. Поэтому очень важно разработать удобную, с понятным и максимально простым интерфейсом систему, которая позволит максимально быстро найти лучшие предложения по интересующей модели автомобиля.

Следовательно, грамотно разработанное программное средство упростит и ускорит процесс выбора и аренды автомобиля, а менеджерам позволит наиболее тесно и эффективно общаться с потенциальными клиентами.

Информационные технологии оказали большое влияние на автомобильную промышленность в целом.

Например, ИТ-подразделения достигли значительных успехов в инфраструктуре самих транспортных средств, включая контроль выбросов, развлечения, функции безопасности и многое другое. Кроме того, ИТ повлияли на отношения между клиентами и арендодателями, в сфере аренды автомобилей [1].

1 Сегодня основной источник информации об автомобилях доступных для аренды – это веб-сайты. Из тех, кто на рынке арендует автомобиль, более половины посещают сайты компаний арендодателей, многие разговаривают с семьей и друзьями, а другие посещают компании лично.

2 Даже при интенсивном посещении сайтов все еще есть потенциал для дальнейшего роста за счет увеличения объема предоставляемого контента.

3 Пользователи сайтов в режиме онлайн упрощают сравнение транспортных средств. Они в том числе хотели бы видеть и отчеты о состоянии [2].

4 Таким образом, целью работы является совершенствование процесса аренды автомобильной посредством его автоматизации на основе web-приложения.

1 Объектом исследования является сфера услуг аренды автомобилей.

2 Предмет исследования – методы и основные задачи автоматизированной поддержки процесса аренды автомобилей, принципы программной реализации систем поддержки процесса аренды автомобилей.

### **3 Актуальность.**

4 Данное программное средство будет являться источником актуальной и полезной информации для потенциальных арендаторов автомобилей, а также платформой для размещения объявлений об имеющихся автомобилях для менеджеров.

5 Система предоставит возможность снизить затраты менеджеров на поиск и привлечение клиентов, а также улучшит производительность менеджеров, что позволит обрабатывать большее количество заявок одному сотруднику.

6 Для пользователя разработанная система улучшит подход к отслеживанию статуса заявки и ускорит процесс поиска информации о прокатных автомобилях.

Следовательно, актуальность данного дипломного проекта бесспорна: грамотно разработанное программное средство упростит и ускорит процесс выбора и аренды автомобиля, а менеджерам позволит наиболее тесно и эффективно общаться с потенциальными клиентами.

### **Выбор и обоснование технических и программных средств для реализации.**

Для достижения поставленной цели потребовалось изучить рынок аренды автомобилей, проанализировать существующие на рынке аналоги и определить их преимущества и недостатки, спроектировать базу данных, разработать функциональное наполнение приложения и пользовательский интерфейс программного продукта, выполнить тестирование информационной системы.

К основным бизнес-требованиям данного сайта для аренды автомобилей относятся:

- возможность поиска, фильтрации и сортировки данных по необходимым критериям и условиям. Это позволит ускорить обработку данных, а в следствии повысить производительность и уменьшить время ожидания ответа;

- ускорить обработку запроса на аренду;

- пользователям должна быть предоставлена возможность обратной связи [3].

Поставленные задачи были решены с помощью следующих ресурсов:

- для построения функциональной модели разрабатываемого продукта использовалось приложение ErWin Process Modeler r7;

- для построения спецификации вариантов использования и моделей представления системы использовалось приложение Enterprise Architect 12.3.

7 Разрабатываемый программный модуль представлен серверной и клиентской частями.

8 Серверная часть реализована в виде REST-сервиса, написанного на языке программирования Java, к которому для получения данных будет обращаться клиентское web-приложение.

9 Для связывания сущностей базы данных с классами Java решено использовать ORM-технология Hibernate.

Написание серверной части программы выполнялось в интегрированной среде разработки IntelliJ IDEA 2017.3, а клиентской – в JetBrains WebStorm 2017.1.

В ходе написания приложения также были использованы следующие паттерны проектирования: MVC (Model-view-controller) и Repository (Репозиторий).

Model-view-controller (MVC) – это паттерн проектирования, с помощью которого модель данных приложения, пользовательский интерфейс и взаимодействие с пользователем разделены на три отдельных компонента таким образом, чтобы модификация одного из компонентов оказывала минимальное воздействие на остальные.

MVC часто используется для построения архитектурного каркаса, когда переходят от теории к реализации в конкретной предметной области [4].

Общая структура концепции паттерна MVC для автоматизированной системы представлена на рисунке 1.

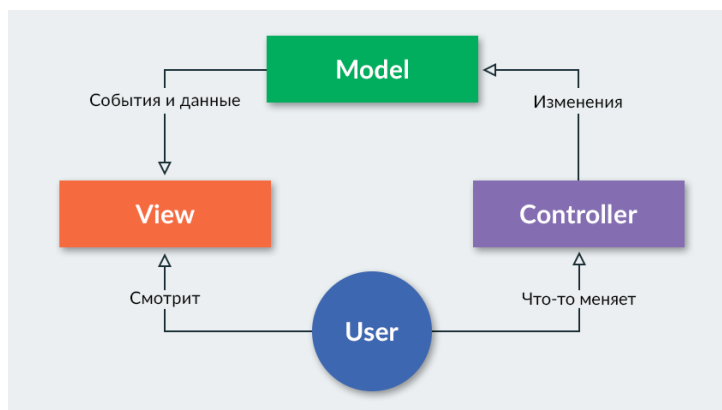


Рисунок 1. Структура работы шаблона MVC

Модель (model) – самая независимая часть системы, она содержит всю бизнес-логику приложения.

Представление (view) отвечает за отображение данных пользователю. Основное предназначение данного компонента – предоставлять информацию из модели в удобном для пользователя формате.

В контроллере (controller) хранится код, который отвечает за обработку действий пользователя. Через контроллер пользователь вносит изменения в данные, хранящиеся в модели.

Рассмотрим подробнее паттерн проектирования Репозиторий (repository). Данный шаблон посредничает между domain-слоем и mapping-слоем, используя интерфейс, схожий с коллекциями для доступа к объектам области определения (рисунок 2).

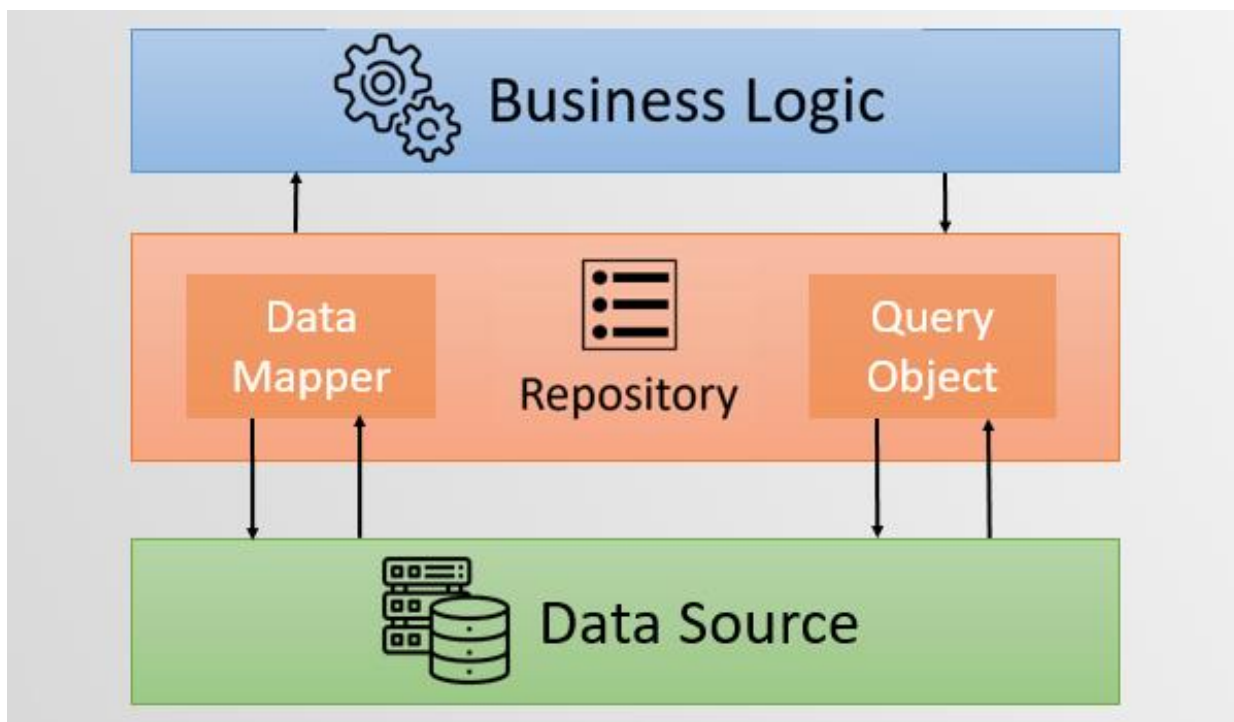


Рисунок 2. Структура работы шаблона Repository

Паттерн Repository посредничает между слоем области определения и слоем распределения данных, работая, как обычная коллекция объектов области определения.

Объекты-клиенты создают описание запроса декларативно и направляют их к объекту-репозиторию (Repository) для обработки.

Объекты могут быть добавлены или удалены из репозитория, как будто они формируют простую коллекцию объектов [5].

Данные паттерны проектирования были использованы в программном приложении с целью повышения гибкости, краткости и понятности программы.

#### **Модели представления программного средства.**

Для подробного описания работы программы будем использовать UML диаграммы.

С помощью схем UML изображаются организационные структуры, моделируются бизнес-процессы (как это уже сделали в диаграмме use cases), проектируются систему на различных этапах и уровнях.

Для лучшего понимания всей системы рассмотрим следующие модели:

- диаграмма вариантов использования;
- диаграмма компонентов;
- диаграмма развертывания.

10 После исследования задач, поставленных перед системой, необходимо перейти к процессу проектирования возможностей предметной области, одним из этапов которого является спецификация вариантов использования системы.

11 Она отражает функциональные возможности и требования системы с использованием действующих лиц и вариантов использования.

12 Варианты использования моделируют службы, задачи, функции, которые должна выполнять система [6].

Актерами, действующими в системе, являются гость, зарегистрированный клиент и менеджер.

Рассмотрим подробнее имеющиеся возможности использования системы. «Пользователю» доступен следующий функционал системы:

– Регистрация. Данная функция включает в себя ввод всех необходимых данных для заполнения профиля.

– Просмотр автопарка. Данный пункт включает в себя просмотр каталога, из которого можно перейти к просмотру всех автомобилей.

– Также пользователь имеет возможность произвести поиск, сортировку или фильтрацию автомобилей по различным параметрам.

13 Зарегистрированный пользователь («Клиент») имеет следующие возможности:

– Обращение в поддержку. В процессе обращения пользователю необходимо будет указать проблему.

– Просмотр новостей. Данный пункт необходим для последующего просмотра заинтересовавшего материала.

– Просмотр профиля. Каждый зарегистрировавшийся пользователь будет иметь возможность просматривать информацию из своего профиля и при необходимости редактировать её.

– Заявка на аренду. Это одна из основных функций системы. Она включает в себя заполнение формы заявки.

14 Менеджеру доступен следующий функционал системы:

– Управление автомобилями.

15 При добавлении авто, например, пользователю необходимо будет ввести информацию об автомобиле: производитель, модель, год выпуска и так далее.

– Управление новостями. Каждый менеджер может добавлять и удалять новости чтобы преподносить клиентам самую актуальную информацию.

– Управление пользователями.



- Предоставляет возможность добавлять новых менеджеров для дальнейшей работы с ними и удалять клиентов.
- Управление контрактами.
- Менеджер может добавлять информацию по заключенным сделкам и просматривать статистику по ним за определённый период времени.

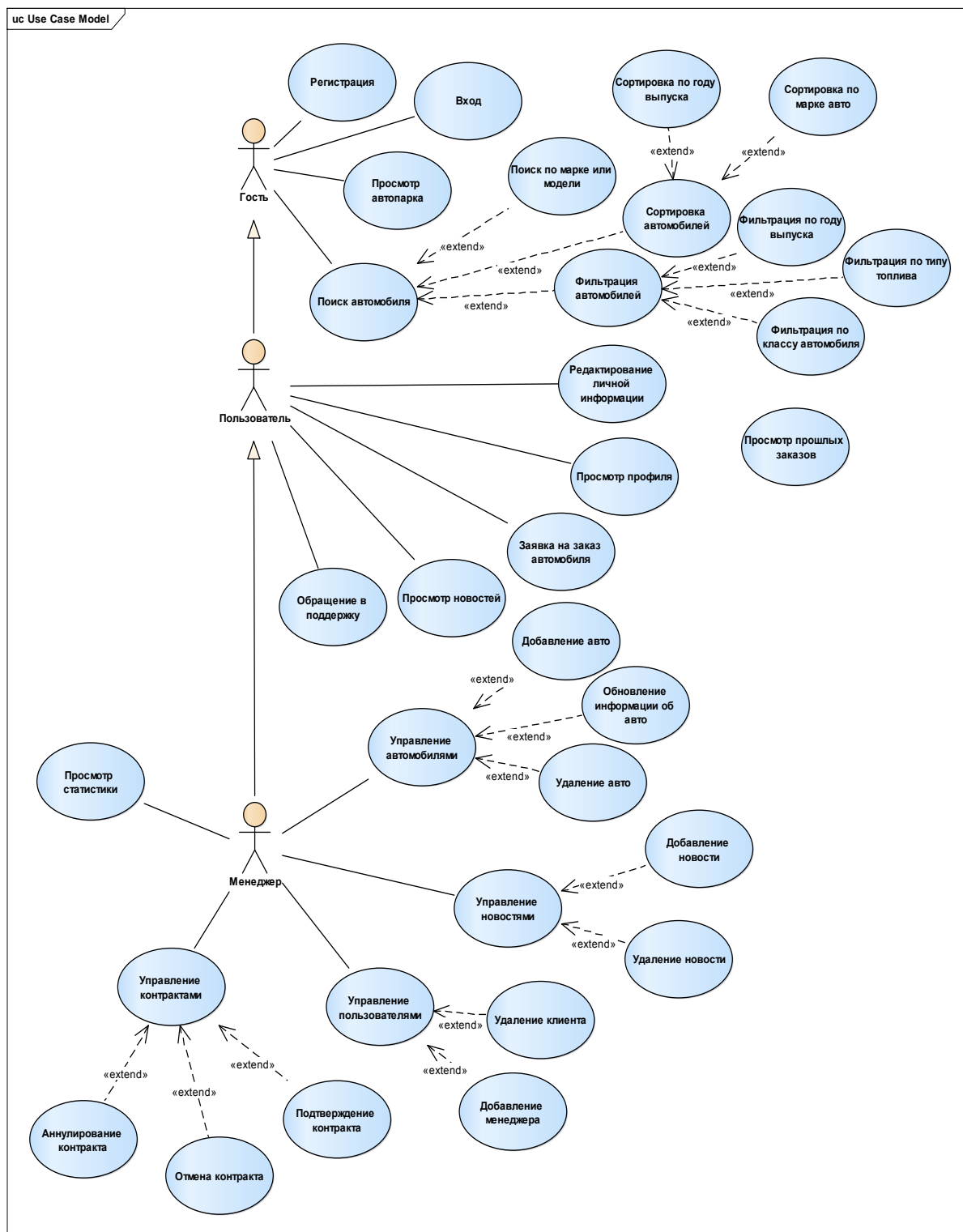


Рисунок 3. Диаграмма вариантов использования мобильного приложения

Для наглядного изображения разбиения программной системы на структурные компоненты и отображения связей (зависимостей) между компонентами были разработаны диаграммы компонентов и развертывания системы, которые представлены на рисунках 4-5.

Диаграмма компонентов – это, по сути, диаграмма классов, которая фокусируется на компонентах системы, которые часто используются для моделирования представления статической реализации системы.

Грамотное использование диаграммы компонентов позволяет:

- 16 – представить физическую структуру компонентов;
- 17 – обратить внимание на компоненты системы и как они связаны;
- акцентировать внимание на поведении системы в части, касающейся интерфейса.

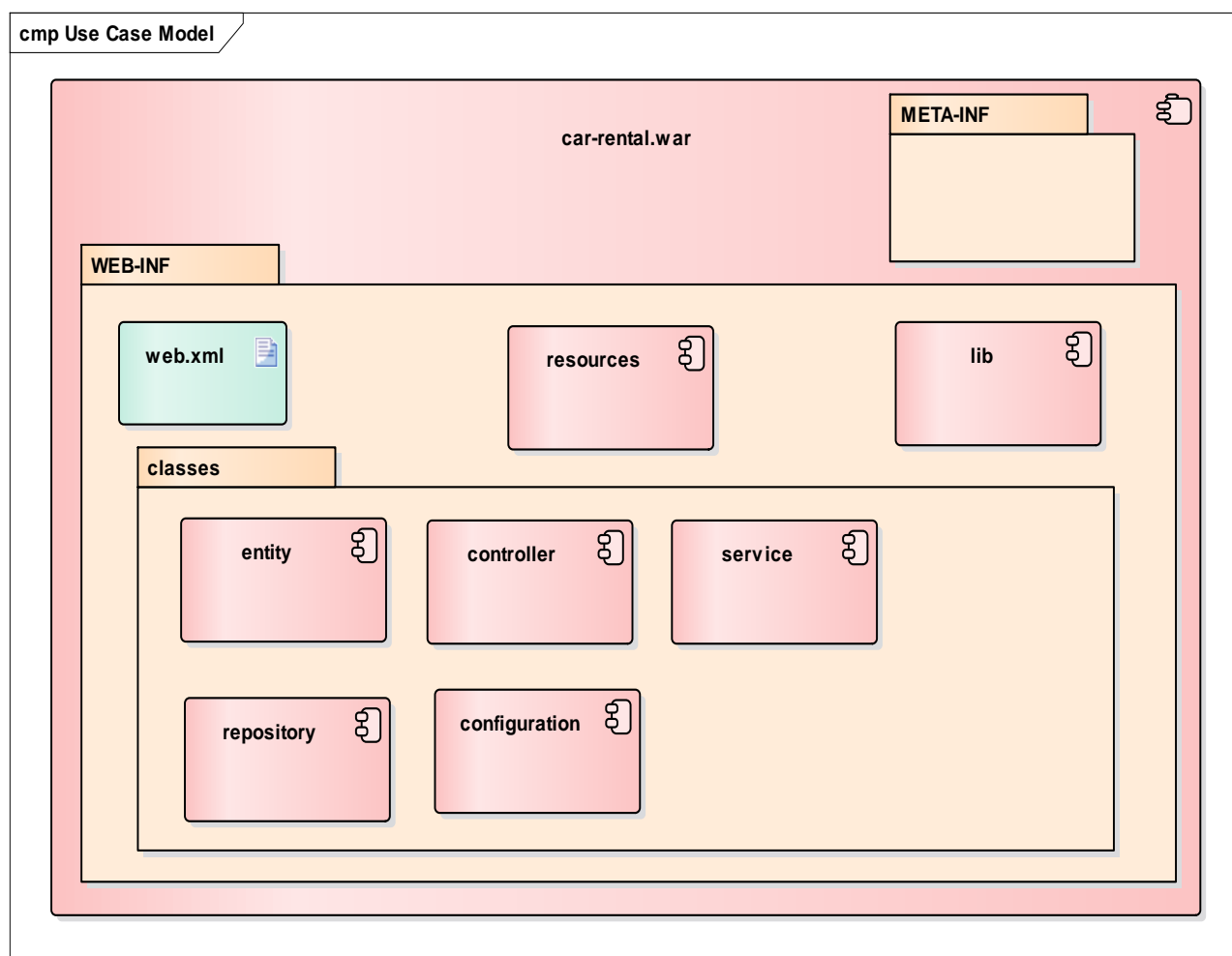


Рисунок 4. Диаграмма компонентов системы

Построение диаграммы компонентов позволяет визуализировать общую структуру исходного кода программной системы, специфицировать исполнимый вариант программной системы, представить концептуальную и физическую схемы баз данных.

В контексте унифицированного языка моделирования (UML) диаграмма развертывания относится к семейству структурных диаграмм, поскольку она описывает аспект самой системы. В этом случае схема развертывания описывает физическое развертывание информации, генерируемой программным обеспечением на аппаратных компонентах [7].

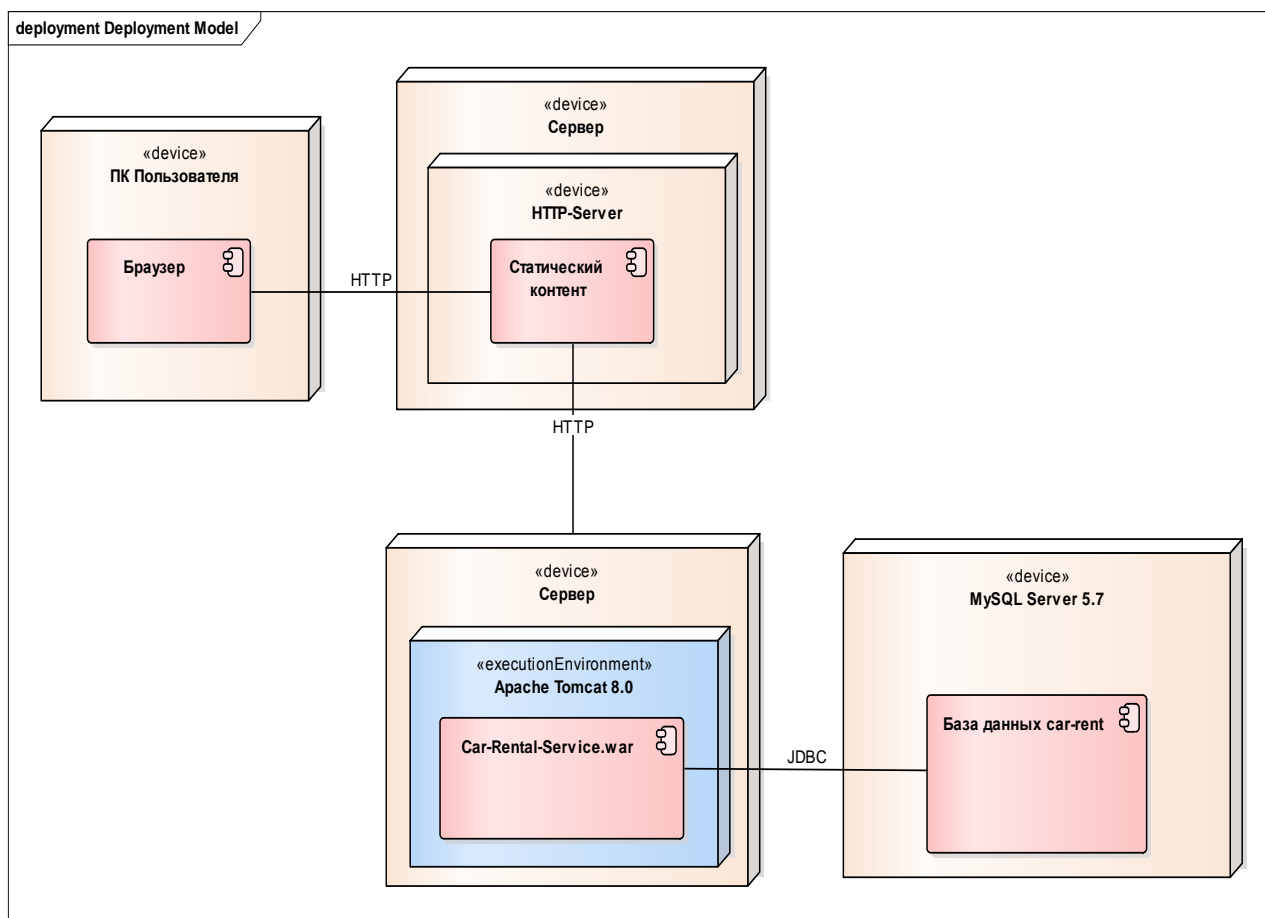


Рисунок 5. Диаграмма развертывания системы

Диаграмма развертывания представляет собой конфигурацию узлов и размещенных в них компонентов.

Диаграмма связана с диаграммой компонентов т. к. в узле размещается один либо несколько компонентов.

Для каждой модели создается только одна диаграмма развертывания, отображающая процессы, устройства и их соединения [7].

#### **Результаты тестирования разработанного ПО.**

Тестирование – это процесс проверки функционала программы с целью подтверждения того, что она работает в соответствии с определёнными требованиями.

Unit-тестирование – это тестирование, которые пишутся, непосредственно, на уровне разработчика (тестирование определённой сущности – метод или класс).

Это крайне важный этап разработки ПО, который помогает создавать качественный продукт.

Для тестирования и проверки работоспособности программного средства были добавлены автоматизированные тесты с использованием фреймворка JUnit.

Главная идея данного фреймворка – сначала тесты, потом код.

Это означает, что сначала мы определяем, что должно получиться в результате работы того или иного куска кода и пишем тесты, которые проверяют идентичность результата с требуемым, после чего пишем сам кусок кода, который и будем тестировать.

Данный подход увеличивает эффективность работы разработчика и позволяет писать более стабильный код. В результате этого мы получаем меньшее количество времени, которое затрачивается на отладку программы.

В процессе разработки системы аренды автомобилей было проведено около двадцати тестов, которые покрывают основной функционал системы. Результаты прохождения тестов представлены ниже (рисунок 6).

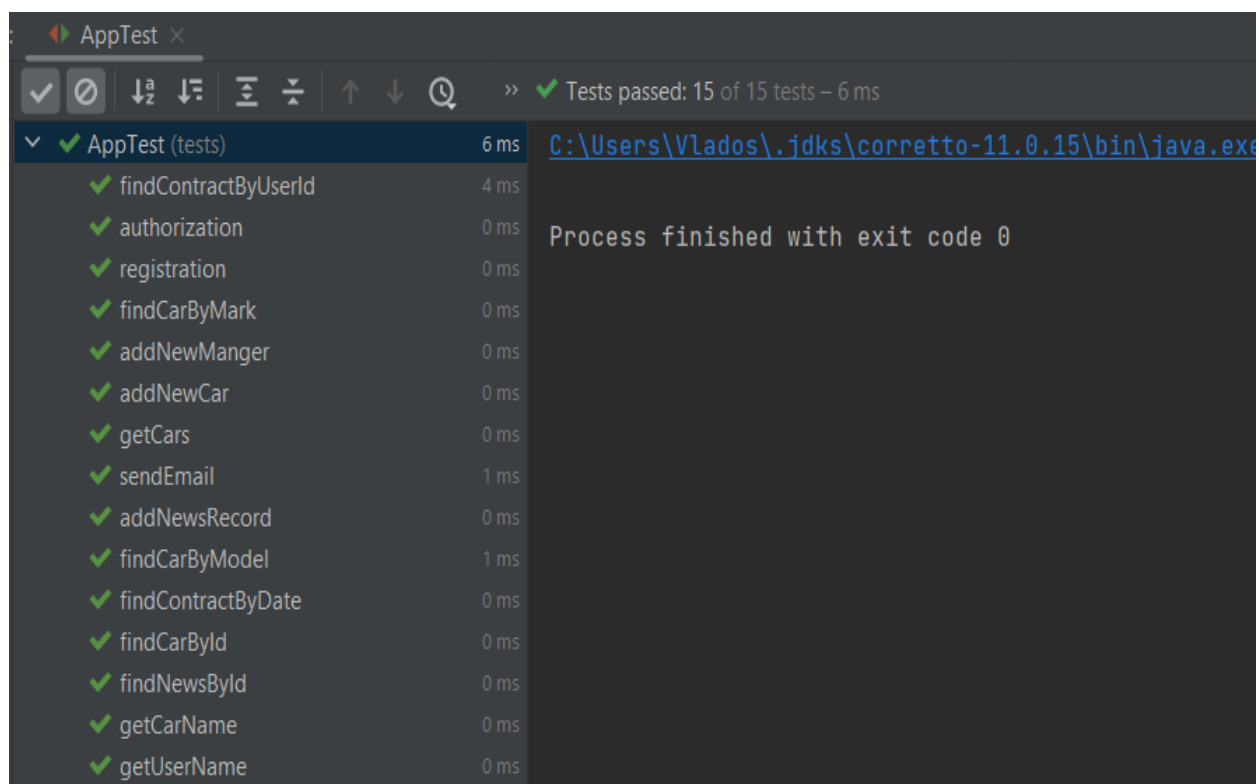


Рисунок 6. Результаты работы тестов

Таким образом, разработанные тесты помогли избежать потери данных и некорректной работы приложения при возникновении исключительных ситуаций.

### **Заключение.**

В процессе работы была изучена специфика взаимодействия менеджеров по аренде автомобилей с клиентами.

Была разработана система, позволяющая менеджерам взаимодействовать с клиентами, а именно делает более удобным добавление автомобилей, редактирование информации по ним и удаление, а также предоставляет возможность хранить информацию о клиентах, добавлять данные по контрактам и просматривать на их основании статистику.

Приложение обладает простым и понятным интерфейсом, так что пользователь любого уровня владения ПК может с ним работать, не испытывая каких-либо трудностей.

В результате применения данной информационной системы будет контролироваться аренда автомобилей между клиентами, сокращено время обслуживания клиентов проката автомобилей, ускорен процесс добавления сведений об аренде автомобилей клиентами, повышено качество предоставления услуг проката автомобилей, сокращены трудоемкость менеджеров и улучшение прочих показателей.

Разработанная программа в полной мере реализует заявленный в ходе проектирования функционал и имеет возможность расширения. В дальнейшем можно расширить функционал путём добавления возможности оплаты аренды автомобиля через приложение с проведением всех необходимых операций.

Подводя итоги, можно сказать, что цель, поставленная при создании проекта, была выполнена: данное приложение повысит качество и скорость обработки заявок, а также увеличит удовлетворенность клиентов, что, в свою очередь, будет привлекать новых арендаторов.

#### **Список литературы**

- [1] Michael G. Albino. International Journal of Science, Engineering and Technology, 2021, 9:2  
[2] Statista: [Электронный ресурс]. URL: <https://www.statista.com/outlook/270/142/car-rentals/>. (Дата обращения: 10.04.2023).  
[3] Julsrud, T.E.; Farstad, E. Car sharing and transformations in households travel patterns. Energy Res. Soc. Sci. 2020, 66  
[4] Описание паттерна MVC: [Электронный ресурс]. URL: <https://javarush.ru/groups/posts/2536-chastjh-7-znakomstvo-s-patternom-mvc-model-view-controller> (Дата обращения: 15.04.2023).  
[5] Описание паттерна Repository: [Электронный ресурс]. URL: <http://design-pattern.ru/patterns/repository.html>. (Дата обращения: 16.04.2023).  
[6] Буч, Г. Язык UML. Руководство пользователя. 2-е изд. / Г. Буч, Д. Рамбо, И. Якобсон. – М. : ДМК Пресс, 2006. – 496с.  
[7] Гома, Х. UML-проектирование / Х.Гома. – М. : ДМУ Пресс, 2011. – 704 с.

### **SOFTWARE SUPPORT FOR CUSTOMER INTERACTION MANAGEMENT IN THE CAR RENTAL INDUSTRY**

**V.A Milentev**  
*Master student of  
Department of Engineering  
Psychology and  
Ergonomics, assistant of the  
department of economic  
informatics BSUIR*

**A.M. Prudnik**  
*Associate Professor,  
Department of Engineering  
Psychology and  
Ergonomics, BSUIR,  
Candidate of Technical  
Sciences, Associate  
Professor*

**I.V.Andryalovich**  
*Deputy Dean of the  
Faculty of Computer  
Design of BSUIR,  
postgraduate student  
of the Department of  
IPIE*

**A.V. Vorobey**  
*Postgraduate student  
of the Department of  
IPIE, Assistant of the  
Department of IPIE,  
Master of  
Engineering*

*Department of Engineering Psychology and Ergonomics  
Faculty of Computer Engineering  
Belarusian State University of Computer Science and Radio Electronics, Republic of Belarus  
E-mail: milvova864@gmail.com*

**Annotation.** This paper describes a web application for a car rental company. This software tool is a source of current and useful information for potential tenants, as well as a platform for placing ads for car rental managers.

**Keywords:** car rental, manager, client, web-application, Java, MySQL.

УДК 004.021:004.75

## ИГРА ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ «КОСМОНАВТ БЕСКОНЕЧНОСТИ»



**М.В. Большакова**

Учащаяся Учреждения образования «Национальный детский технопарк», учащаяся ГУО «Лицей г. Новополоцка»



**М.С. Ильясова**

Магистрант кафедры инженерной психологии и эргономики БГУИР



**Ф.В. Усенко.**

Магистрант кафедры инженерной психологии и эргономики БГУИР



**Л.Р. Коркин**

Магистр технических наук, ассистент кафедры инженерной психологии и эргономики БГУИР



**А.М. Прудник**

Доцент кафедры инженерной психологии и эргономики БГУИР, кандидат технических наук, доцент  
[aleksander.prudnik@bsuir.by](mailto:aleksander.prudnik@bsuir.by)

### **М.В. Большакова**

Обучается в лицее г. Новополоцка. Область научных интересов связана с технологиями виртуальной и дополненной реальности.

### **М.С. Ильясова**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с автоматизированным тестированием информационных систем.

### **Ф.В. Усенко**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой манипуляторов для дистанционного управления.

### **Л.Р. Коркин**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с системами распознавания снимков.

### **А.М. Прудник**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с взаимодействием человека с компьютером, интерфейсами информационных систем, пользовательскими интерфейсами, front-end web development.

**Аннотация.** Рассматриваются преимущества и актуальность разработки VR приложений на примере проекта «Космонавт бесконечности». Также большое внимание уделяется описанию сценарию игры и некоторым аспектам разработки приложения с приведением примеров скриптов и наглядных изображений

отдельного функционала. Помимо этого, проект был протестирован и готов в качестве бета-версии, а также было предоставлено описание целевой аудитории приложения.

**Ключевые слова:** VR-технологии, головоломки, виртуальная реальность, космос, Unity, C#.

### **Введение.**

Индустрия игр виртуальной реальности в настоящий момент бурно развивается, однако на данный момент сложно сказать, что игры виртуальной реальности однозначно лучше, чем классические игры на десктопных компьютерах. И те, и другие имеют свои преимущества и недостатки. Тем не менее, кажется, что разработка VR-игр имеет гораздо больший потенциал с точки зрения погружения в среду и реалистичного игрового процесса.

Поэтому, как только отрасль решит некоторые проблемы, с которыми она все еще сталкивается, она, скорее всего, займет большую часть рынка [1].

Одно из преимуществ VR-игр по сравнению с классическими видеоиграми заключается в том, что они предоставляют иммерсивный опыт.

Обзор окружающего пространства в 360° обеспечивает реалистичность, с которой тяжело конкурировать плоским дисплеям.

Более того, ношение очков виртуальной реальности позволяет пользователям меньше отвлекаться на свое физическое окружение. Если же к этому добавить реалистичную механику взаимодействия и постоянно улучшающуюся графику, то получается действительно захватывающий опыт [2].

Дополнительным преимуществом игр виртуальной реальности является их требование к движению игроков, что полезнее, чем пребывание в статическом положении.

Наконец, VR-игры дают лучший многопользовательский опыт – пользователь чувствует, что его спутники находятся с ним в одном пространстве.

Целью проекта являлась разработка виртуальной игры «Космонавт бесконечности» для демонстрации возможностей VR-технологий и предоставление уникального пространства для ощущения новых эмоций.

Актуальность разработки игр виртуальной реальности подтверждается в том числе и статистическими данными.

Так, например, предполагается, что потребительские расходы на VR и AR к 2024 году достигнут 72 млрд USD, а общее количество продаваемых VR и AR гаджетов уже сейчас исчисляется десятками миллионов в год [3].

### **Постановка задачи и выбор инструментов.**

Для достижения поставленной цели необходимо было решить следующие задачи:

1. Выбрать жанр и продумать место действий в данной игре;
2. Провести обзор аналогов разрабатываемой системы и выполнить сравнительный анализ их достоинств и ограничений;
3. Прописать сюжет для главного героя и игры;
4. Спроектировать модели персонажей и сцены
5. Реализовать механику движения и тактильного взаимодействия с игровыми объектами;
6. Провести тестирование VR-игры.

В качестве среды разработки нами была выбрана кроссплатформенная среда разработки игр Unity, а в качестве языка программирования C#.

### **Разработка сценария игры.**

Для реализации игры была выбрана последовательность жанров как головоломки для стимулирования и развития умственных способностей и поддержание здорового ума любого человека.



Данная стратегия направлена на совершенствование и тренировку мышления путем предоставления игрокам интерактивных и когнитивно-стимулирующих заданий, что является важным аспектом в области развития человеческого интеллекта.

В начале игры игрок попадает из космического корабля на островок новой планеты, где в процессе ему придётся восстанавливать поврежденный корабль при помощи несложных головоломок.

Игрок может взаимодействовать с отдельными компонентами корабля для активации уровня с мини-игрой.

В игре присутствуют разные головоломки, некоторые из них представлены ниже.

Головоломка «Квадрат» – игра на логику и интуицию.

Игровое поле представляет собой квадрат с ячейками, в которые вписаны цифры (рисунок 1). Необходимо зачеркнуть все цифры в квадрате.

Цифры обозначают количество ячеек, которые будут зачеркнуты при нажатии.

Игрок должен начинать двигаться от знака S, при этом за границы квадрата выходить нельзя, зачеркивание по диагонали, также не допускается. Пример прохождения первого уровня по шагам: 1 вниз, 1 вправо, 2 вверх, 1 вправо, 2 вниз.

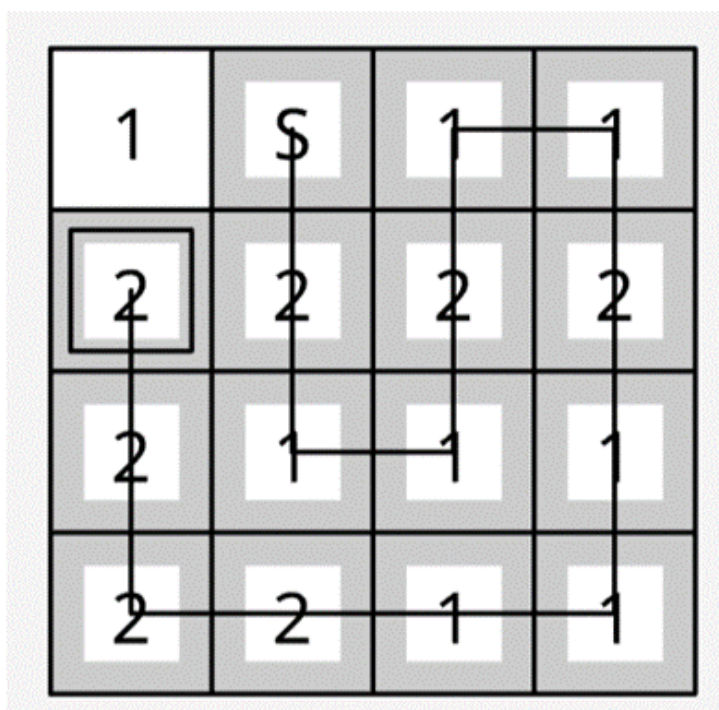


Рисунок 1. Графическое представление головоломки «Квадрат»

Головоломка со сдвигом – головоломка, имеющая некое сходство с тетрисом (рисунок 2). Сходство заключается в концепции игрового поля, похожего на «стакан» игры тетрис, и в механике смещения рядов из блоков.

Здесь заполненная линия также исчезает, принося игроку очки.

Однако фигуры являются обыкновенными прямоугольными блоками разных цветов и длин. Блоки добавляются рядами (линиями) снизу стакана. Переполнение последнего, как и в тетрисе, ведет к окончанию игры.

Задачей играющего является сдвиг выбранной фигуры вправо или влево с целью повлиять на заполнение свободных участков на любом из рядов игрового поля. Если ряд заполнен, образуется линия, которая тут же исчезает.



Если цвет всех брусков, составляющих линию, одинаков, то игрок получает повышенное число очков.

Гравитация в игре правильная, поэтому лежащие сверху блоки осыпаются вниз и становятся на свободные места. За один ход можно подвинуть только один брусок.

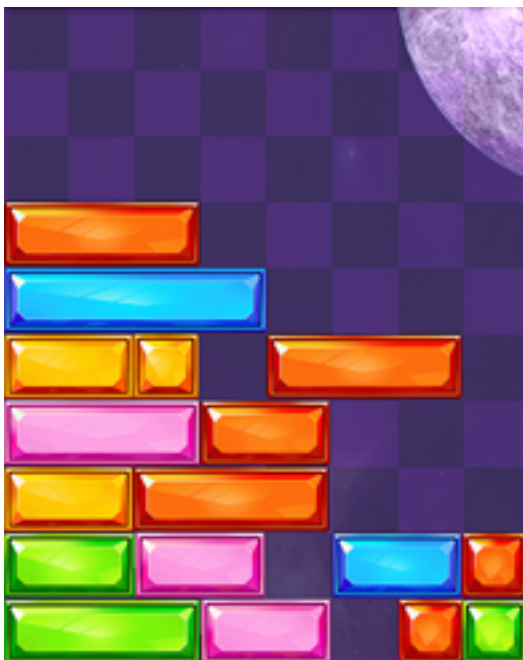


Рисунок 2. Наглядный вид головоломки со сдвигом

Головоломка «Линии» – относительно несложная головоломка, в которой нужно соединять пары точек одинакового цвета (рисунок 3), при этом линии соединения не должны пересекаться.



Рисунок 3. Вид головоломки «Линии»

Для выполнения данной задачи игрок должен выбрать цветную точку и, удерживая ее, нарисовать линию от этой точки до соответствующей точки другого цвета. В процессе игры указанный маршрут будет преобразовываться в линию, соответствующую цвету

точек. Фрагменты линий могут располагаться как горизонтально, так и вертикально. Решение головоломки требует заполнения каждой клетки на поле, при этом некоторые клетки уже заполнены цветными точками, в то время как остальные заполняются фрагментами линий в процессе игры. Для успешного решения головоломки необходимо убедиться в заполнении всех клеток поля [4].

#### Реализация головоломок игры.

Для реализации головоломки «Квадрат» был написан скрипт, который будет обрабатывать пользовательский ввод и отображать текущее состояние игрового поля. При запуске скрипт загружает первый уровень игры, который будет представлять собой квадрат с несколькими цифрами внутри ячеек.

Пользовательский ввод будет состоять из команд, которые будут задавать направление движения и количество ячеек, которые нужно зачеркнуть.

Скрипт проверяет корректность введенных данных благодаря методу `IsValidMove()`, представленном на рисунке 4, и обрабатывает команды соответствующим образом.

Например, если пользователь вводит команду "1 вниз", то скрипт перемещает игрока на одну ячейку вниз и зачеркнуть соответствующее количество ячеек.

При каждом изменении состояния игрового поля отображается текущее состояние на экране и проверяется, были ли зачеркнуты все цифры в квадрате. Если все цифры были зачеркнуты, то скрипт загружает следующий уровень игры.

```
static bool IsValidMove(int[,] board, int row, int col)
{
    if (row < 0 || row >= board.GetLength(0) ||
        col < 0 || col >= board.GetLength(1))
        return false;
    if (board[row, col] == -1)
        return false;
    return true;
}
```

Рисунок 4. Скрипт, отвечающий за корректность введенных данных игроком

Для реализации Головоломки со сдвигом было разработано игровое поле, которое было определено параметрами, такими как размер и количество строк и столбцов. Далее были созданы фигуры, состоящие из прямоугольных блоков разных цветов и длин, которые будут добавляться в игровое поле рядами сверху. При сдвиге фигуры вправо или влево, необходимо проводить проверки. Для этого был разработан скрипт, использующий метод `MoveFigure()`, который позволяет сдвигать фигуру на заданное расстояние. Пример реализации данного скрипта представлен на рисунке 5.

```
public void MoveFigure(int xShift)
{
    if ((currentCol + xShift < 0) || (currentCol + xShift + figureWidth >
boardWidth))
        return;
    ClearFigure(currentRow, currentCol);
    currentCol += xShift;
    DrawFigure(currentRow, currentCol);
}
```

Рисунок 5. Скрипт, реализующий сдвиг фигуры

Для реализации головоломки "Линии" был создан скрипт на языке программирования C#, который позволяет игроку выбирать цветную точку и соединять ее линией с соответствующей точкой другого цвета, при этом удерживая ее на экране. Для выполнения данной задачи, скрипт проверяет цвета точек и учитывает линии соединения, чтобы они не пересекались. Скрипт был создан с использованием различных методов и функций, таких как проверка цвета точек, проверка пересечений линий соединения и их соединение в правильном порядке. В процессе игры, скрипт преобразует маршрут линий соединения в линию, соответствующую цвету точек, которые игрок соединил. Фрагменты линий могут располагаться как горизонтально, так и вертикально, что добавляет сложности и интереса в игру. Таким образом, класс Node содержит информацию о вершине графа, заданной через позицию в двумерном пространстве, а также связи с другими вершинами и ребрами графа. Кроме того, класс Node определяет логику взаимодействия вершин и ребер между собой, что влияет на игровой процесс. Каждая вершина может быть раскрашена в один из цветов, которые хранятся в GameManager. Две вершины могут быть соединены только если они имеют одинаковый цвет. Метод UpdateInput() определяет поведение вершины при ее соединении с другой вершиной. Метод AddEdge(connectedNode), представленный на рисунке 6, добавляет ребро между текущим узлом и переданным узлом connectedNode в головоломке "Линии". Метод устанавливает цвет connectedNode таким же, как и у текущего узла, добавляет connectedNode в список связанных узлов текущего узла и добавляет текущий узел в список связанных узлов connectedNode. Затем метод находит игровой объект, соответствующий соединению между текущим узлом и connectedNode, и устанавливает его активным. Далее, метод устанавливает цвет соединения между текущим узлом и connectedNode, используя цвет, связанный с цветом узла.

```
private void AddEdge(Node connectedNode)
{
    connectedNode.colorId = colorId;
    connectedNode.ConnectedNodes.Add(this);
    ConnectedNodes.Add(connectedNode);
    GameObject connectedEdge =
ConnectedEdges[connectedNode];
    connectedEdge.SetActive(true);

connectedEdge.GetComponent<SpriteRenderer>().color =
    GameManager.Instance.NodeColors[colorId %
GameManager.Instance.NodeColors.Count];
}
```

*Рисунок 6. Скрипт, отвечающий за добавление ребра*

Для обеспечения качественной и стабильной работы приложения, были проведены тесты. В процессе тестирования были найдены и устранены ошибки, которые приводили к неправильной работе приложения. После завершения тестирования игровой продукт был готов к использованию в качестве бета-версии, так как остались некоторые недочеты.

#### **Результаты и обсуждение.**

В процессе достижения цели были решены следующие задачи: 1. Выбран жанр и продумано место действий в данной игре; 2. Проведён обзор аналогов разрабатываемой системы и выполнен сравнительный анализ их достоинств и ограничений; 3. Прописан сюжет для главного героя и игры; 4. Спроектированы модели персонажей и сцены

5. Реализованы механики движения и тактильного взаимодействия с игровыми объектами;
6. Проведено тестирование VR-игры.

### **Заключение.**

В результате разработки получилось реализовать основные задуманные механики и представить бета-версию VR игры-головоломки «Космонавт бесконечности». Этот продукт может быть нацелен на любую возрастную категорию, но в основном подойдёт для людей в возрасте от 8 до 50+ лет, также в данную игру может играть человек любого вероисповедания и национальности.

### **Список литературы**

- [1] Компьютерные игры как средство вхождения в виртуальную реальность [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/kompyuternye-igry-kak-sredstvo-vhozhdeniya-v-virtualnuyu-realnost/viewer>. (Дата обращения: 25.04.2023).
- [2] Основные преимущества игр в виртуальной реальности [Электронный ресурс]. URL: <https://fainaidea.com/jeto-interesno-znat/osnovnye-preimushhestva-igr-v-virtualnoj-realnosti-191607.html>. (Дата обращения: 25.04.2023).
- [3] Гейтс Билл. Дорога в будущее. — М.: Изд. отд. «Рус. ред.» ТОО «Channel Trading Ltd.», 1996. — 312 с ISBN 5-7502-0019-1.
- [4] Головоломка линии [Электронный ресурс]. URL: <https://www.min2win.ru/game/golovolomka-linii.html>. (Дата обращения: 25.04.2023).

## **VIRTUAL REALITY GAME ASTRONAUT INFINITY**

***M.V. Bolshakova***

*Student of the educational  
Institution "National  
Children's Technopark",  
student  
GUO "Lyceum of  
Novopolotsk"*

***M.S. Piyasova***

*Master's student of the  
Department of Engineering  
Psychology and Ergonomics of  
BSUIR*

***P.V. Usenko***

*Master's student of the  
Department of Engineering  
Psychology and Ergonomics of  
BSUIR*

***L.R. Korkin***

*Master of Technical Sciences,  
Assistant of the Department of  
Engineering Psychology and  
Ergonomics of BSUIR*

***A.M. Prudnik***

*Associate Professor of Engineering  
Psychology and Ergonomics  
Department of BSUIR, Candidate of  
Technical Sciences, Associate  
Professor*

*Department of Engineering Psychology and Ergonomics  
Faculty of Computer-Aided Design  
Belarusian State University of Informatics and Radio Electronics, Republic of Belarus  
E-mail: [aleksander.prudnik@bsuir.by](mailto:aleksander.prudnik@bsuir.by)*

**Abstract.** The advantages and relevance of the development of VR applications are considered on the example of the project "Cosmonaut of Infinity". Also, much attention is paid to the description of the game scenario and some aspects of application development with examples of scripts and visual images of individual functionality. In addition, the project was tested and ready as a beta version, and a description of the target audience of the application was provided.

**Keywords:** VR technologies, brain-teaser, virtual reality, space, Unity, C#.

УДК 625

## ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ МЕЖКОМПОНЕНТНЫХ ФАЗОВЫХ ОТНОШЕНИЙ



**Д.А. Кечик**

Ассистент кафедры информационных радиотехнологий БГУИР, магистр технических наук  
[ya.dan.kechik@yandex.by](mailto:ya.dan.kechik@yandex.by)



**И.Г. Давыдов**

Доцент кафедры информационных радиотехнологий БГУИР, кандидат технических наук, доцент  
[davydov\\_ig@bsuir.by](mailto:davydov_ig@bsuir.by)

### **Д.А. Кечик**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с построением алгоритмов обработки сигналов вибрации с целью диагностики оборудования.

### **И.Г. Давыдов**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с построением алгоритмов обработки сигналов вибрации с целью диагностики оборудования.

**Аннотация.** В настоящей работе представлена аппроксимация распределения межкомпонентных фазовых отношений (МКФО) составляющих полигармонического сигнала зашумлённого гауссовским шумом. Показано, что для высоких отношений сигнал-шум закон распределения МКФО близок к нормальному.

**Ключевые слова:** межкомпонентные фазовые отношения, цифровая обработка сигналов, статистика, тест Колмогорова-Смирнова.

### **Введение.**

Один из способов существенного повышения эффективности производства – внедрение обслуживания оборудования по его фактическому состоянию. Состояние работающего агрегата оценивается методами неразрушающего контроля, в том числе методами виброакустической диагностики.

Затраты производства снижаются за счёт уменьшения времени простоя оборудования, предотвращения его внезапных отказов, уменьшения числа ошибок монтажа при сборке за счёт удлинения межремонтного интервала, своевременного выявления подобных ошибок [1].

Одной из распространённых проблем сборки оборудования, приводящая к значительному росту частоты отказа его узлов, является расцентровка соединяемых валов [2]. Выделяют два вида расцентровок: угловая и параллельная.

В первом случае, оси соединяемых валов представляют собой пересекающиеся прямые, во втором – параллельные.

Наличие обоих видов расцентровки приводит к скрещивающимся осям валов.

В работе [3] предложено использовать МКФО для различения вида и степени расцентровки валов. Показано, что среднее значение МКФО зависит от величины их смещения, а эмпирическое распределение МКФО близко к нормальному.

Цель данной работы – получить теоретическое распределение некоторых МКФО, построить зависимость СКО МКФО от отношения сигнал-шум (ОСШ) и длины реализации. Это поможет в дальнейшем построить алгоритм принятия решения о виде и степени дефекта, оценить его параметры.

### Статистические распределения фазы.

Представим вибрационный сигнал диагностируемого оборудования в виде:

$$x(t) = \sum_{h=1}^H A(h) \cos(2\pi h F_0 t + \phi(h)) + n(t) = \sum_{h=1}^H A(h) \cos(2\pi h F_0 t + \hat{\phi}(h)), \quad (1)$$

$$\phi(h) = h\phi_0 + \theta(h) = E[\hat{\phi}(h)]$$

где  $H$  – число квазигармонических колебаний,  $A(h)$  – амплитуда  $h$ -го колебания,  $\phi_0$  – начальная постоянная фаза компонентов вибрации, обусловленная случайным моментом начала записи.  $\theta(h)$  – постоянная величина, обусловленная запаздыванием возникновения кинематических сил, порождающих колебания на разных частотах. Значения  $A(h)$  и  $\theta(h)$  могут быть получены путём решения системы дифференциальных уравнений, составленных с учётом кинематической схемы дефектного узла. Здесь эти значения считаются априори известными, как и дисперсия шума  $\sigma$ .

Рассмотрим распределение спектральных отсчётов  $X(hF_0)$ ,  $h = 1, 2, \dots, H$ :

$$X(f) = \sum_{t=0}^{T-1} x(t) \cdot \exp(-2\pi j f \cdot t). \quad (2)$$

В силу линейности преобразования Фурье (ПФ),  $X(hF_0) = A(h) + N(hF_0)$ , где  $N(hF_0)$  – ПФ шума:

$$N(f) = \sum_{t=0}^{T-1} n(t) \cdot \exp(-2\pi j f \cdot t). \quad (3)$$

Согласно [4], мнимые и действительные части спектральных отсчётов  $N(f)$  при условии стационарности  $n(t)$  распределены асимптотически по нормальному закону:

$$D(\Re[N(f)]) = D(\Im[N(f)]) = N [0, T \cdot P(f)] \quad (4)$$

где  $T$  – дискретное время,  $P(f)$  – спектральная плотность мощности, определяемая как:

$$P(f) = \sum_{\tau=-\infty}^{\infty} \exp(2\pi j f \tau) \cdot c(\tau) \quad (5)$$

$$c(\tau) = \text{cov}[X(t+\tau), X(t)], \text{cov}(X, Y) = E[X - E(X)] \overline{E[Y - E(Y)]} \quad (6)$$

Т.к.  $n(t)$  – белый гауссов шум, то  $c(\tau) = \sigma^2 \delta(0)$ , соответственно,  $P(f) = P_0 = \sigma^2 / F_s$ , где  $\sigma^2$  – мощность шума,  $F_s$  – частота дискретизации, тогда:

$$D(\Re[N(f)]) = D(\Im[N(f)]) = N [0, T \cdot P_0] \quad (7)$$

Если  $X(f)$  нормировать по числу отсчётов  $T$ , дисперсия Фурье-отсчётов уменьшится в  $T^2$  раз:

$$D(\Re[N(f)]) = D(\Im[N(f)]) = N [0, \sigma_s^2], \sigma_s^2 = T^{-1} \cdot P_0 \quad (8)$$

Распределение аргумента комплексной величины  $X(hF_0)$  рассчитывается через совместное распределение мнимой и действительной частей [6]:

$$D(\Re[X(hF_0)], \Im[X(hF_0)]) = (2\pi\sigma_s^2)^{-1} \cdot \exp[-(\Re[X(hF_0)] - a)^2 + (\Im[X(hF_0)] - b)^2] / (2\sigma_s), a = A(h)\cos(\phi(h)), b = A(h)\sin(\phi(h)) \quad (9)$$

которое после преобразования декартовых координат в полярные с учётом

$$\begin{aligned} \Re[X(hF_0)] &= r(h)\cos\hat{\phi}(h), \\ \Im[X(hF_0)] &= r(h)\sin\hat{\phi}(h), \end{aligned} \quad (10)$$

где  $r, \nu$  – амплитуда и фаза Фурье-отсчёта  $h$ -й гармоники соответственно, преобразуется в совместное распределение амплитуды и фазы:

$$D(r(h), \hat{\phi}(h)) = r(2\pi\sigma_s^2)^{-1} \cdot \exp[-(r(h)\cos\hat{\phi}(h))^2 + (r(h)\sin\hat{\phi}(h) - b)^2] / (2\sigma_s) \quad (11)$$

откуда распределение фазы получено интегрированием (11) по амплитуде:

$$D(\hat{\phi}(h)) = \int_0^\infty D(r(h), \hat{\phi}(h)) dr = \frac{\exp(-A^2(h)/2\sigma_s^2)}{2\pi} + \frac{A\cos u}{2\sigma_s\sqrt{2\pi}} \cdot \left[ 1 + \operatorname{erf}\left(\frac{A\cos u}{\sqrt{2\sigma_s}}\right) \right] \cdot \exp\left(-\frac{A^2}{2\sigma_s^2}\sin^2 u\right), u = \hat{\phi}(h) - \phi(h) \quad (12)$$

Согласно [5], при условии  $(A/\sigma_s)\cos(\phi - \hat{\phi}(h)) > 3$  возможно аппроксимировать распределение (12):

$$D(\hat{\phi}(h)) \approx \frac{A}{\sigma_s\sqrt{2\pi}} \cdot \cos u \cdot \exp\left[-\frac{A^2}{2\sigma_s^2}\sin^2 u\right]. \quad (13)$$

При условии малого отклонения фазы этот закон приближается нормальным распределением с СКО  $\sigma_\phi(h)$  и средним  $\phi(h)$ :

$$D(\hat{\phi}(h)) \approx \frac{A(h)}{\sigma_{phi}\sqrt{2\pi}} \cdot \exp\left[-\frac{A(h)^2}{2\sigma_{phi}^2}(\hat{\phi}(h) - \phi(h))^2\right], \quad (14)$$

$$\sigma_\phi(h) = \sigma_s / A(h) = P_0 / (\sqrt{T} \cdot A(h)). \quad (15)$$

### Статистические распределения МКФО.

Существуют линейные комбинации фаз составляющих сигнала (1), не зависящие от момента начала отсчёта  $t_0$ .

Такие комбинации называются межкомпонентными фазовыми отношениями (МКФО). Запишем общее выражение МКФО:

$$\begin{aligned}\Theta(k) &= \sum_{h=0}^H k(h) \cdot \hat{\phi}(h) = \sum_{h=1}^H k(h) \cdot (h\phi_0 + \theta(h) + \gamma(h)) = \\ &= \phi_0 \cdot \sum_{h=1}^H k(h)h + \sum_{h=1}^H k(h) \cdot (\theta(h) + \gamma(h))\end{aligned}\quad (16)$$

Здесь  $k \in Z$  – коэффициенты линейной комбинации,  $\gamma(h) = \hat{\phi}(h) - \phi(h)$ . Выражение (16) не зависит от  $\phi_0(h) = 2\pi \cdot hF_0 \cdot t_0$  при условии

$$\sum_{h=1}^H k(h)h = 0 \quad (17)$$

Известно, что распределение суммы  $Z$  независимых нормальных случайных величин  $X, Y$  также нормальное с дисперсией, равной сумме дисперсий исходных величин:  $D(Z) = D(X) + D(Y)$ . Кроме того,  $D(C \cdot X) = C^2 D(X)$ . Отсюда распределение  $\Theta(k)$ :

$$D(\Theta(k)) = \frac{1}{\sigma_{\Theta}(k)\sqrt{2\pi}} \cdot \exp\left[-\frac{(\Theta(k) - \Theta_0(k))^2}{2\sigma_s^2}\right], \quad (18)$$

$$\Theta_0(k) = \sum_{h=1}^H k(h) \cdot \theta(h), \quad (19)$$

при этом СКО  $\Theta(k)$  выражается через длину реализации, значения амплитуд гармоник и спектральной плотности мощности шума:

$$\sigma_{\Theta}(k) = \sqrt{\sum_{h=1}^H \sigma_{\phi}^2(h)} = \sqrt{\sum_{h=1}^H P_0^2 / (T \cdot A(h))^2} = \sqrt{P_0 / (T) \cdot \sum_{h=1}^H A^{-2}(h)} \quad (20).$$

### Результаты.

Эмпирические распределения МКФО были получены путём моделирования. Сигналы вида (1) генерировались при заданных значениях амплитуд, фаз, уровня шума,  $H = 3$ .

Эмпирические и теоретические распределения  $\hat{\phi}(h), \Theta(k)$  при  $k = \{2, -1, 0\}$  и  $k = \{1, 1, -1\}$  приведены на рисунках 1-3 а-в соответственно. СКО теоретического распределения  $\phi(h)$  рассчитывалось по формуле (15), распределения  $\Theta(k)$  – по формуле (20).

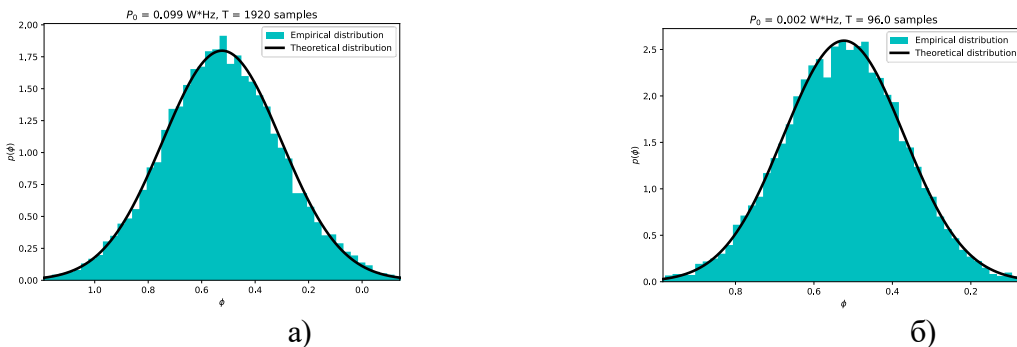


Рисунок 1. Распределения начальной фазы первой гармоники: а)  $P_0 = 0.099, T = 1920$ ; б)  $P_0 = 0.002, T = 96$



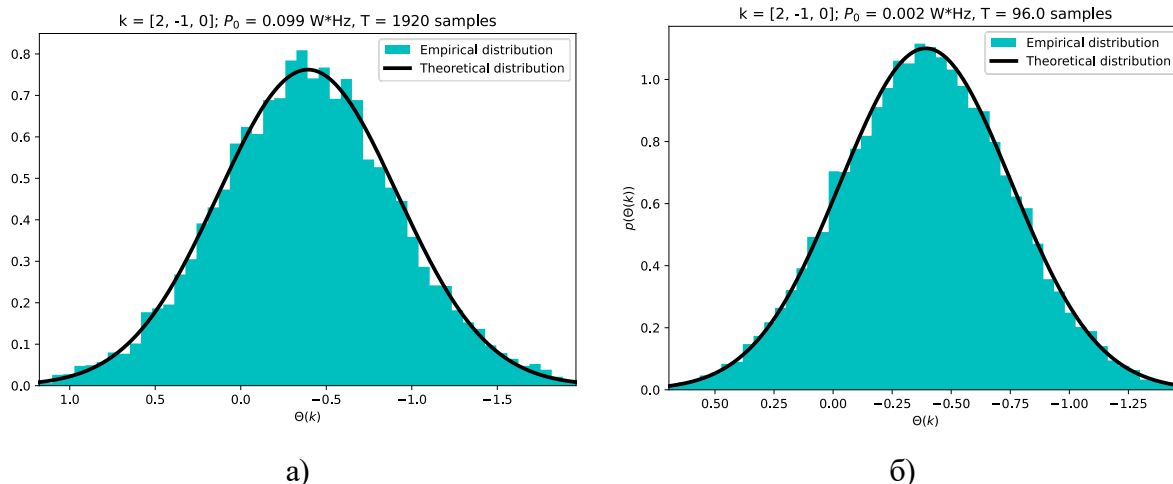


Рисунок 2. Распределения  $\Theta(k)$ ,  $k = [2, -1, 0]$ : а)  $P_0 = 0.099$ ,  $T = 1920$ ; б)  $P_0 = 0.002$ ,  $T = 96$

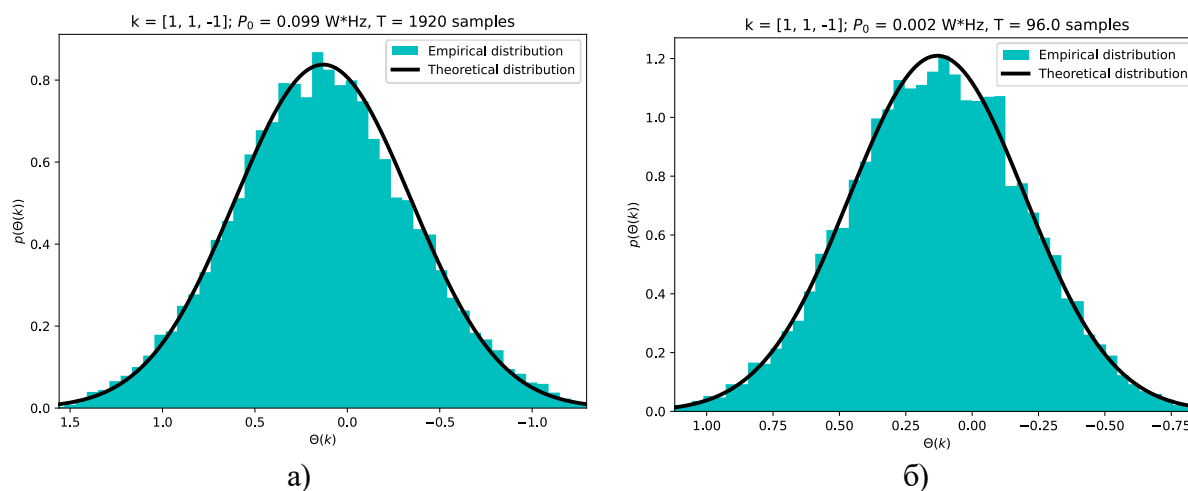


Рисунок 3. Распределения  $\Theta(k)$ ,  $k = [1, 1, -1]$ : а)  $P_0 = 0.099$ ,  $T = 1920$ ; б)  $P_0 = 0.002$ ,  $T = 96$

Соответствие эмпирического и теоретического распределений проверялась при помощи теста Колмогорова-Смирнова. Суть теста заключается в оценке статистики

$$S_n = \sup_{|x| < \infty} |\hat{F}(x) - F(x)|, \quad (21)$$

где  $\hat{F}(x)$  – эмпирическая функция распределения,  $F(x)$  – теоретическая.

Нулевая гипотеза не отвергается, если выполняется неравенство  $p = P(S > S_n) > \alpha$ , где  $S$  – критическое значение статистики для уровня значимости  $\alpha$ . Нулевая гипотеза была сформулирована как:

- $D(\hat{\phi}(h)) = N[\phi(h), \sigma_\phi^2(h)]$
- $D(\Theta(k)) = N[\Theta_0(k), \sigma_\Theta^2(k)]$

Проверялась нулевая гипотеза при  $A = \{1, 0.8, 0.7\}$ ,  $F_0 = 10$  Гц,  $F_s = 960$  Гц.

Проводилось два теста: при постоянной СПМ  $P_0=0.002$  Вт/Гц и  $T = [96...960]$  отсчётов,  $T = 1920$ ,  $P_0 = [0.002...0.2]$  Вт/Гц. При спектральной плотности мощности шума  $P_0 \leq 0.1$  Вт/Гц и  $T=1920$   $p \geq 0.05$ , что не позволяет отвергнуть  $H_0$  в этих случаях.

Полученный результат свидетельствует о том, что при высоких ОСШ распределение МКФО можно считать нормальным со средним (19) и СКО (20).

### **Заключение.**

Межкомпонентные фазовые отношения, вычисленные для Фурье-коэффициентов сигнала вибрации, при высоких ОСШ распределены по нормальному закону.

Соответствие эмпирического распределения теоретическому закону проверялась тестом Колмогорова-Смирнова.

Получены аналитические выражения расчёта среднего и дисперсии МКФО, учитывающие длину реализации, значения амплитуд и уровень шума.

### **Список литературы**

- [1]. Абрамов, И.Л. Вибродиагностика энергетического оборудования / И.Л. Абрамов. – Кемерово : КузГТУ, 2011. – 81 с.
- [2]. Hariharan, Dr.V. Vibration analysis of misaligned shaft – ball bearing system / Dr.V. Hariharan // Indian J. Sci. Technol. – 2009. – Т. 2. – С. 45-50.
- Кечик, Д.А. Различение неисправностей оборудования при помощи межкомпонентной фазовой обработки сигналов / Д.А. Кечик, И.Г. Давыдов. – 2022.
- [3]. Brillinger, D.R. Time Series: Data Analysis and Theory / D.R. Brillinger. – Society for Industrial and Applied Mathematics, 2001. – 564 p.
- [4]. Левин, Б.Р. Теоретические основы статистической радиотехники / Б.Р. Левин. – Изд. 3. – М. : Радио и связь, 1989.
- [5]. Bennett, W. Methods of Solving Noise Problems / W. Bennett // Proc. IRE. – 1956. – Vol. 44, № 5. – P. 609-638.

## **PROBABILITY DENSITIES OF INTER-COMPONENT PHASE RELATIONS**

**Д.А. Кечик**

Ассистент кафедры информационных  
радиотехнологий БГУИР, магистр технических  
наук  
ya.dan.kechik@yandex.by

**И.Г. Давыдов**

Доцент кафедры информационных  
радиотехнологий БГУИР, кандидат технических  
наук, доцент  
davydov\_ig@bsuir.by

Department of Information Radiotechnologies  
Faculty of Radiotechnics and Electronics  
Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: ya.dan.kechik@yandex.by

**Abstract.** Approximation of inter-component phase relations (ICPR) of components of corrupted by additive gaussian noise polyharmonic signal has been presented in this paper. ICPR probability distribution law has been shown to be close to normal.

**Key words:** inter-component phase relations, digital signal processing, statistics, Kolmogorov-Smirnov test.

УДК 621.311

## АЛГОРИТМЫ ПОИСКА НЕИСПРАВНОСТЕЙ СОЛНЕЧНОЙ ЭЛЕКТРОСТАНЦИИ НА ОСНОВЕ БОЛЬШИХ ДАННЫХ ЕЁ МОНИТОРИНГА



**К.С. Дик**

*Аспирант Белорусского государственного университета информатики и радиоэлектроники*



**И.И. Пилецкий**

*Доцент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники, к. ф.-м. н.*

### **К.С. Дик**

*Аспирант БГУИР. Инженер по программному обеспечению и данным в Ютех Солжунс. Проводит научные исследования по обнаружению аномалий с помощью искусственной нейронной сети.*

### **И.И. Пилецкий**

*Доцент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники. В сфере IT более 50 лет. Участие в разработке нескольких десятков крупных проектов: главный конструктор проекта, главный архитектор программного и информационного обеспечения, руководитель проекта, начальник отдела.*

**Аннотация.** Представлен анализ алгоритмов поиска неисправностей солнечных электростанций на основе больших данных собранных в результате мониторинга с использованием облачных ресурсов. Проанализированы различные подходы к моделированию солнечных панелей и солнечных ячеек, исследованы системы управления с обратной связью и влияние изменений параметров окружающей среды на производительность солнечных электростанций. Особое внимание уделено сравнению различных моделей солнечных ячеек для определения влияния входных параметров на их эффективность.

**Ключевые слова:** обнаружение аномалий, солнечная PV панель, глубокое обучение, неконтролируемое обучение, математические модели.

### **Введение**

Солнечные электростанции являются важным источником возобновляемой энергии, и повышение эффективности их работы является актуальной задачей. В данной статье рассматриваются различные подходы к оптимизации работы солнечных панелей в сложных и нелинейных условиях с использованием информационных систем.

Целью исследования явился анализ существующих методик, подходов и алгоритмов поиска неисправностей солнечных электростанций на основе больших данных, которые собираются в результате сбора телеметрии по каждой из солнечных панелей электростанции, а также поиск возможности оптимизации эффективности работы солнечных панелей.

### **Мониторинг и диагностика состояния солнечных панелей**

В отношении солнечных панелей наиболее актуальным применением информационных систем является оптимизация эффективности их работы в различных, нелинейных и сложно прогнозируемых условиях.

Для мониторинга и диагностики состояния панелей, а также прогнозирования их дальнейшего состояния чаще всего применяется моделирование. При помощи моделирования возможна как начальная оценка производительности при установке солнечных панелей, так и долгосрочное планирование за счет изменения параметров,

влияющих на производительность панели. Также при расхождении значений на текущей реальной панели со значениями модели можно вовремя обнаружить неисправность и устранить ее.

В работе автора Heng Wang [1] моделируются и оцениваются не только сами солнечные панели, но и станция сотовой связи, которая будет потреблять полученное электричество (рис. 1).

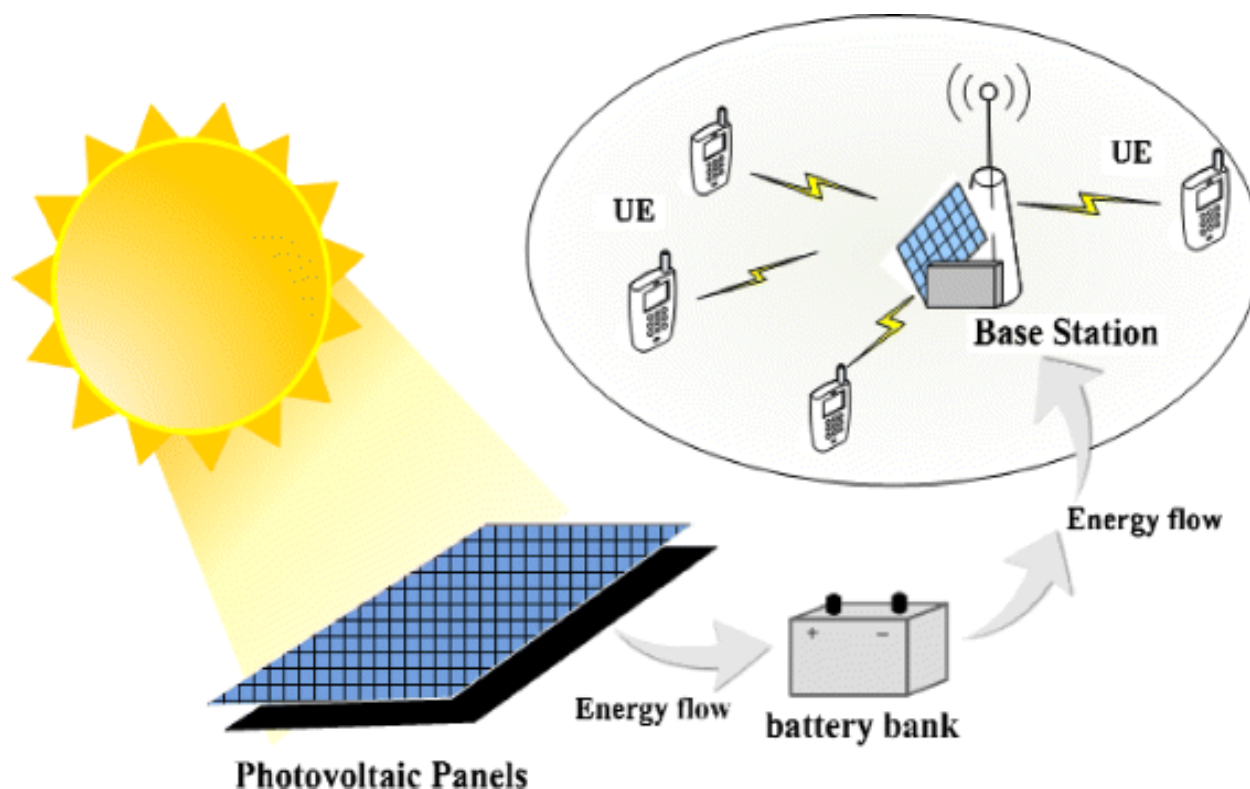


Рисунок 1. Схема работы базовой станции сотовой связи с использованием солнечных панелей

Помимо стандартных вводятся специальные метрики для оценки эффективности преобразования солнечной энергии, возможных перебоев в работе системы и средней глубиной разряда накопленной энергии. С использованием метрик создается оптимальная модель, решающая в том числе и проблему размеров станции, заключающейся в балансе между надежностью и компактностью. Кроме различных метрик для оценки были также использованы метеорологические данные по солнечному излучению и температуре.

В статье [2] исследуется влияние толщины стекла и поливинилфторида, температуры на входе системы, температуры ячейки и других параметров на эффективность солнечных панелей. Рассматривается возможность оптимизации эффективности по температуре и эффективности по электричеству при помощи генетических алгоритмов. В данных алгоритмах используются все параметры, на которые можно влиять (например интенсивность солнечного излучения таким параметром не является).

Авторы [3] проводили моделирование не для отдельной панели, а для массива панелей, учитывая нелинейность его характеристик и постоянно меняющиеся условия работы. Для этого представлены модели ячеек/панелей/массивов панелей с использованием Tag tools в Matlab/Simulink. Солнечная панель DS-100M использовалась как reference model (рис. 2).



Рисунок 2. Установка солнечной панели DS-100M

При моделировании учитываются как изменяющиеся условия окружающей среды, так и физические параметры панелей. В результате моделирования была получена точная, надежная и легко модифицируемая модель массива солнечных PV панелей, включающая изменяющиеся значения температуры, солнечного излучения, частичного затенения и физических параметров панелей.

В работе [4] рассматривается моделирование солнечных ячеек и определение параметров солнечной ячейки в реальном времени. Проводится сравнение эффективности моделирования между моделью электрической цепи с 1 диодом и с 2-мя диодами с использованием численных методов. Данная статья нацелена на выбор наиболее подходящей модели солнечной ячейки после моделирования.

В [5] исследованы солнечные ячейки различной геометрической формы (не стандартные плоскостные, а с изогнутыми поверхностями, рис. 3). В работе предложены модели для создания более гибких ячеек оптимальной формы, позволяющие использовать солнечные панели на крышах зданий нетипичной формы. Сотни различных вариаций размещения и соответствующие им результирующие характеристики солнечных панелей были изучены. Также в процессе использовались панели различных размеров и эффективность панелей сравнивалась с аналогичными панелями плоскостной формы.

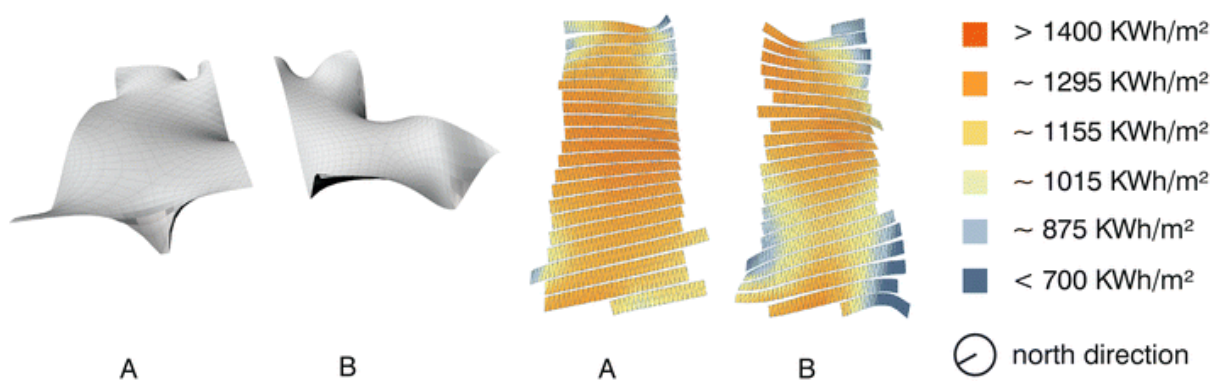


Рисунок 3. Примеры различных форм панелей для крыш и сравнение распространения света в зависимости от формы

Авторы [6] рассматривают математическую модель солнечной панели для оценки МРР. Учитываются отклонения данных и сопротивление шунта, для расчета динамического изменения которых используется метод Ньютона. Модель успешно протестирована на контроллере dSPACE DS1104.

В работе [7] использованы нейронные сети для прогнозирования мощности, производимой монокристаллической PV панелью. В процессе эксперимента были измерены и зафиксированы значения тока и напряжения, производимые панелью в течение года. Также использовались значения солнечного излучения, температуры воздуха, скорость ветра, направление ветра, относительная влажность и угол наклона Солнца. В результате была получена модель, в которой RMS (root mean square, среднеквадратичная ошибка) не превышала 1,4% и коэффициент корреляции R варьировался от 99,637 до 99,998% на тестовом наборе данных. Натренированная в данной работе нейронная сеть позволяет оценивать панели на любой местности с использованием атмосферных показателей. Полученная нейронная сеть предназначена для первоначального прогнозирования эффективности PV панелей при различных природных условиях в месте предполагаемой установки панели.

Статья [8] рассматривает системы управления с обратной связью и зависимость небольших изменений мощности в PV системах при изменениях параметров окружающей среды (температура и солнечное излучение). Данные изменения ведут себя как возмущения в системе и должны быть точно оценены в виде изменений в результирующих мощности и/или токе. Предложена линейная модель, которая принимает на вход температуру и солнечное излучение. На выходе получаем ток и мощность массива солнечных ячеек. Данная модель протестирована на панелях в лаборатории при различных условиях параметров окружающей среды.

Авторы исследования [9] предложили модель солнечной панели в виде модели эквивалентной ей электрической цепи для оценки электрических параметров панели. Главная проблема заключалась в точности данных параметров. В модели использовалось 7 параметров, результаты модели были протестированы на 6 панелях, созданных с применением различных технологий. Среди них присутствовала монокристаллическая, поликристаллическая и модель с тонкой пленкой. Результаты модели совпали с результатами оценки I-V кривых на данных панелях. Также результаты полученной модели сравнивались с другими уже существующими моделями, использующими 5 и 7 параметров. По результатам полученная модель оказалась более точной по сравнению с аналогами.

В работе [10] рассмотрены различные модели солнечной ячейки: на основе одного диода, на основе одного диода с  $R_s$  (последовательным соединением), на основе одного диода с  $R_p$  (параллельным соединением), на основе двух диодов и на основе трех диодов. Для определения влияния входных параметров на эффективность ячейки использовалась SPICE симуляция. Получены следующие результаты: фототок  $I_L$  пропорционален солнечному излучению,  $R_s$  уменьшает напряжение короткого замыкания и коэффициент заполнения.  $R_p$  сокращает напряжение разомкнутой цепи (холостое напряжение) и оба диода (диффузионный и рекомбинационный) сокращают значения напряжения разомкнутой цепи и коэффициента заполнения. Также показано, что увеличение температуры ячейки снижает значения напряжения разомкнутой цепи и коэффициента заполнения, и поэтому значительно снижает производительность ячейки.

Авторы [11] исследовали точность двухдиодной модели для трех различных технологий PV ячеек: монокристаллической, некристаллической (аморфной), и технологии на основе микроморфного кремния (рис. 4). При проверке модели использовались I-V и P-V кривые при стандартных STC. Сравнивались результаты, полученные при помощи модели и результаты с реальных солнечных панелей данных типов, полученные в



солнечный и в облачный дни. Найдены подходящие коэффициенты для модели при различных уровнях солнечного излучения и температуры.

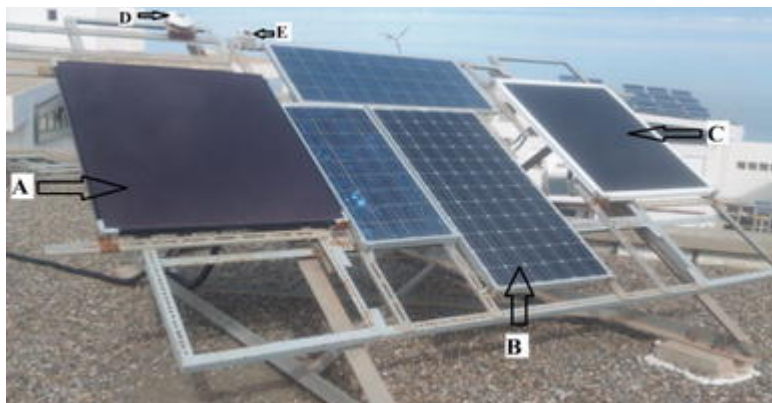


Рисунок 4. Различные виды технологий производства PV панелей: А - кремниевая аморфная, В - кремниевая монокристаллическая, С - кремниевая микроморфная. D и E - пиранометр и кремниевая reference cell соответственно

В работе [12] рассмотрены текущие проблемы численного моделирования кристаллических кремниевых солнечных панелей. Предложенная модель позволяет ускорить оптимизацию произведенных ячеек и прогнозировать изменения в ячейках. Предлагается использовать параметры для обычных моделей солнечных панелей и приспособить их конкретно для солнечных панелей, например внутреннюю плотность носителей, подвижность неосновных носителей, рекомбинацию пассивированных плоскостей и оптические модели. Текущими основными затруднениями при проектировании солнечных панелей названы подвижность дырок неосновных носителей, рекомбинация же при низких плотностях легирующей примеси и промежуточные уровни впырка и тонко настроенные параметры панели в зависимости от температур.

В статье [13] предложена модель солнечной ячейки, позволяющая точно прогнозировать I-V характеристики при различных значениях солнечного излучения и температуры. Модель значительно оптимизирует вычисления для модели с одним диодом и 5-ю параметрами, вводя 5 граничных условий, основанных на данных, полученных от производителя панели. Вычисления сведены к нахождению 2-х неизвестных параметров -  $R_s$  (series resistance) и  $R_{sh}$  (shunt resistance), которые сходятся в среднем за 5 итераций. Параметры модели экстраполированы для учета изменений температуры и солнечного излучения. Данные сходятся с коммерческими PV панелями с высокой точностью. RMS для предложенной I-V модели равен 0,0031, что превосходит аналогичные модели.

В работе [14] рассмотрена проблема различных природных условий в зависимости от места размещения солнечных панелей. Изучалось нахождение оптимальных природных условий для новых панелей с максимальной эффективностью. За основу были взяты STC (солнечное излучение 1000 Вт/м<sup>2</sup>, температура 25 °C и коэффициент воздушных масс 1,5. Данные условия являются идеальными, но практически не встречаются на практике из-за климатических особенностей, так что в лучшем случае несколько параметров будут иметь близкие к STC значения.

Авторы [15] рассматривают модель для оптимизации массивов PV панелей при помощи самоорганизующейся конфигурации. Предложен метод на основе генетических алгоритмов, максимизирующий выходную мощность массива и снижающий количество переключений. Рассмотрена система из 16 панелей (по 4 в каждом соединении, последовательные и параллельные). Также используется дополнительный метод

оптимизации производительности при различных уровнях затенения. Результаты показывают улучшения в производительности от 106.49 до 171.03% по сравнению со статичным вариантом использования массивов. Отдельно стоит отметить оптимизацию скорости поиска наиболее подходящей конфигурации: вместо  $1616 = 1,84 * 1019$  вариаций в худшем случае необходимо рассмотреть лишь  $132 * 50 = 6600$ . Последнее особенно актуально для работы в реальных условиях и со значительно большим количеством панелей, расчет оптимальной конфигурации для которых должен проходить в реальном времени. Также исключены лишние переключения панелей (если для оптимального состояния не нужно переключение и панель уже находится в нужном положении). Для работы данной системы необходимо лишь наличие двух базовых датчиков: один для тока и другой для напряжения на массиве панелей.

Работа [16] рассматривает модели с 5-ю параметрами, включающие управляющий ток генератора и сопротивления диода, шунта и последовательное сопротивление. Модель основана на данных, полученных в режимах разомкнутой цепи, максимальной мощности и короткого замыкания. Также в модели учитывается частичное затенение в случае двух параллельно соединенных аморфных солнечных панелей. Полученные при помощи данной модели данные были провалидированы при помощи экспериментальных измерений на реальных PV панелях.

В работе [17] рассмотрена проблема сложности оценки эффективности солнечных панелей на стадии проектирования, преимущественно, монтирующихся на крышах зданий. Используются следующие параметры:

- 1 горизонтальные компоненты солнечного излучения (глобальное, рассеянное, прямое излучения) на конкретной местности;
- 2 эффекты затенения, преимущественно от других крыш;
- 3 наклон, форма крыши;
- 4 солнечное излучение с учетом наклона крыши;
- 5 свободное место на крыше, предназначенное для установки PV панелей.

Используются физические и эмпирические модели, геостатические методы оценки, метод проб, географические информационные системы (ГИС) и методы на основе определения света и расстояния (LiDAR - light detection and ranging), а также методы машинного обучения. Учитываются плюсы и минусы каждого из данных методов.

Авторы [18] рассматривают влияние затенения на характеристики I - V и P - V кривых солнечной панели. Учитываются температура, солнечное излучение и уровень затенения. Особенно сильно затенение влияет на большие солнечные электростанции. Также из-за затенения вместо одной MPP возникает несколько пиковых значений, что затрудняет анализ эффективности панелей. Предложена модель для оценки I -V и P - V кривых при нестандартном уровне солнечного излучения из-за частичного затенения. Для валидации модели использовались средства симуляции Matlab.

Авторы исследования [29] сравнивают эффективность короткой последовательности массива солнечных панелей с разным типом соединения - последовательным и параллельным. При частичном затенении эффективность массива падает и возникает несколько MPP. Величина потерь на затенении зависит от конфигурации массива, физического расположения панелей и параметров затенения. Использовался небольшой массив PV панелей из 3-х модулей. Измерения и моделирование проводились для различных типов соединения: последовательного и параллельного. Также использовались различные уровни солнечного излучения и параметры затенения. По результатам при последовательном соединении панелей потери значительно превышают аналогичные при параллельном соединении.

В работах [19, 20] рассматривают проблему оценки уровня влияния частичного затенения на эффективность PV панелей. Для данной оценки используется модель, которая



была провалидирована в реальных условиях (рис. 5). Использовались 3 вида затенения, классифицирующиеся по длительности воздействия затенения на панель и по относительному изменению напряжения в точках MPP. Результаты позволяют улучшить будущие методы определения MPP (MPPT).

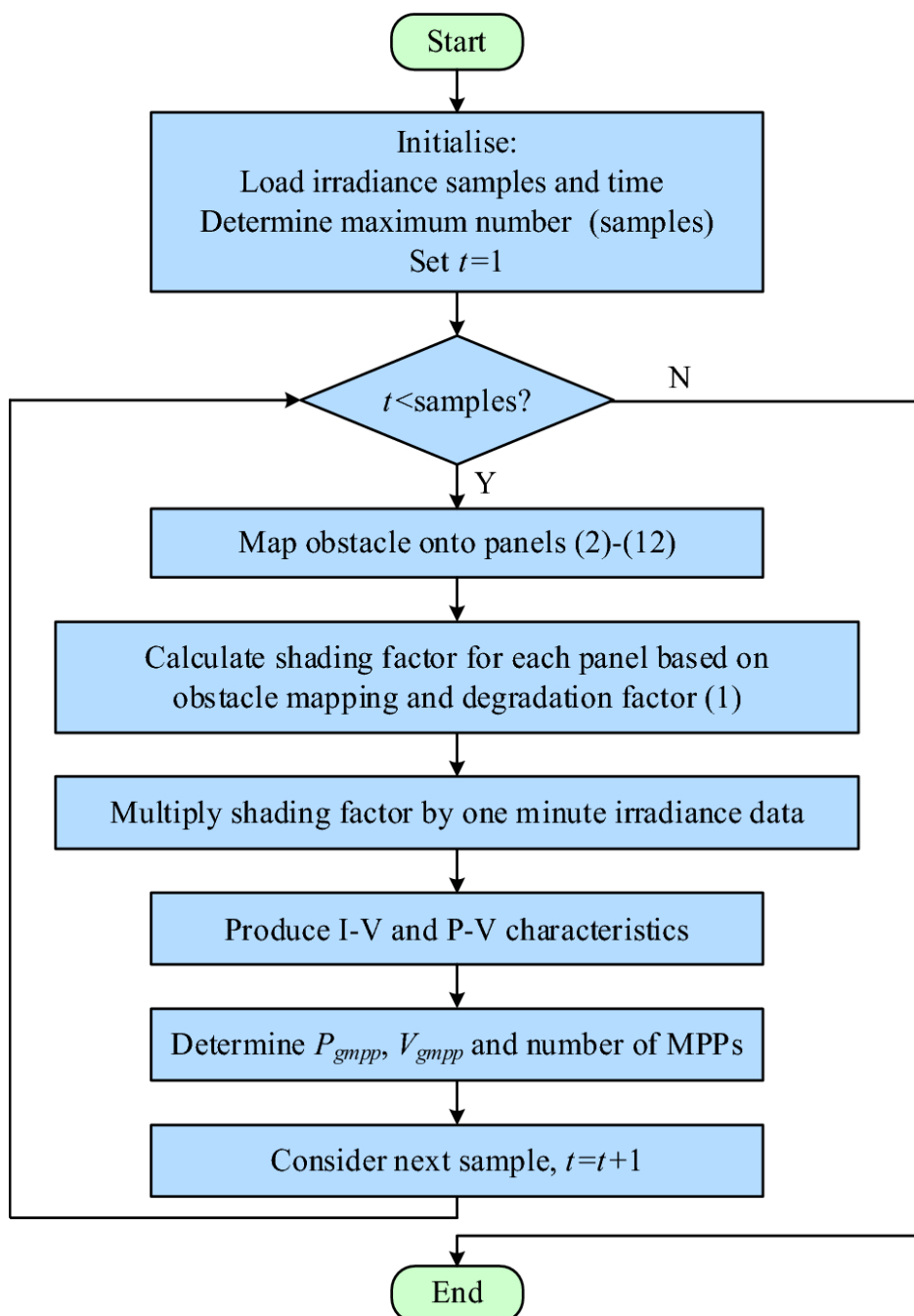


Рисунок 5. Алгоритм моделирования и валидации в условиях частичного затенения

В условиях частичного затенения каждая солнечная ячейка должна моделироваться индивидуально. Предложенная в [21] модель позволяет минимизировать влияние данной особенности на эффективность моделирования. Алгоритм разделяет ячейки в небольшие группы по 2-3 штуки на основе их показателей, представленных в виде коэффициентов.

Коэффициенты рассчитываются и изменяются в реальном времени. Метод представлен и продемонстрирован в Modelica / Dymola.

Авторы [22] пытаются решить проблему точности MPPT алгоритмов в различных и непредсказуемых условиях, а в [23] рассмотрены методы оптимизации и архитектуры для максимизации выработки энергии PV массивом при любых погодных условиях.. В виде решения предлагается создание модели, позволяющей проводить на ней тесты точности MPPT алгоритмов. Данная модель должна полностью включать в себя реализацию MPPT карты, эмулирующей данные с реальной панели. Представлена реализация модели, позволяющей проверить устойчивость результатов MPPT алгоритма при быстро меняющихся условиях среды. Полученные результаты подтверждают надежность и улучшение эффективности при использовании предложенной модели. В [23] Приводится описание актуальных известных методов оптимизации и моделирования PV массивов. Далее оцениваются известные методы MPPT (MPP tracking) и выделяются их основные преимущества и недостатки. Отдельно проводится сравнительный анализ более современных методов с использованием искусственного интеллекта и классических методов отслеживания MPP. Далее рассматриваются возможные применения искусственного интеллекта в других проблемах, связанных с солнечными панелями.

Результаты исследований [24] рассматривают проблему диагностики неисправностей в аморфных PV панелях. Создана модель, нацеленная на оптимизацию электрических показателей и оптимизацию подбора кривой, используемых в MPPT и при диагностике неисправностей. Данные для модели получаются при помощи National Instrument data acquisition system (NI DAQ USB-6212) и применяются при моделировании в Matlab для расчета I-V кривой в реальном времени. Результаты сравниваются с реальными значениями тока и напряжения и проводится диагностика неисправностей. Использовалась трехслойная солнечная панель модели Unit-solar ES-62T.

В [25] рассмотрены мониторинг и диагностика солнечных электростанций. Существуют различные способы мониторинга, к примеру, на некоторых станциях его проводят раз в год, и из-за этого при ухудшениях производительности невозможно определить, чем именно это было вызвано. Чаще диагностика производится в определенные интервалы времени, от минут до нескольких часов. В данном случае возможно своевременное проведение процедуры диагностики неисправностей. В зависимости от системы диагностики определяются постоянные потери энергии, полное отключение, кратковременные потери энергии, затенение. Но в любом случае данные системы не могут определить источник возникших проблем. С появлением систем с использованием DMPPT (distributed MPPT), оптимизаторов мощности и микроинверторы возможности мониторинга и диагностики значительно расширились. Для коммуникации используется модуль WSN (wireless sensor network). Через данные сенсоры передаются следующие данные: напряжение, ток, солнечное излучение и температура. Данные поступают в центр управления, который и производит дальнейшую диагностику. В работе развивается данное направление и в систему MPPT добавляется AC/DC преобразователь, позволяющий определять изменения напряжения в заданном диапазоне, измерять частичную I-V кривую для определения сопротивлений на шунте, диоде, последовательного сопротивления. Используются новые разработанные аналитические и метаэвристические алгоритмы для определения параметров панели. Вычисляемые параметры используются не только для долгосрочной диагностики, но и для своевременного анализа изменений по I-V кривой. Разработана модель для MATLAB, провалидирована с использованием реальных данных.

Авторы [26] осуществили моделирование экономически эффективных PV устройств V-образной формы. LCPV (low concentration photovoltaics) и методы MPPT могут быть интегрированы в PV панели V-образной формы для улучшения их эффективности. Создана модель для симуляции попадания прямого солнечного излучения на V-образную панель.

Данная модель используется для гибкого подбора параметров, высокой детализации геометрии и низких затрат на вычисления. Валидация модели проводилась с использованием непараметрического статистического анализа, который подтвердил высокую точность модели. Также применялся анализ цен и индекс экономической эффективности. Данная модель предназначена для проектирования и сравнения различных V-образных PV панелей.

В [27] приведены результаты моделирования и симуляции солнечных электростанций. Данные этапы необходимы для оценки потенциальной эффективности электростанции после установки в зависимости от местности и климатических условий. Используется однодиодная эквивалентная модель цепи. Рассматриваются I-V и P-V кривые. Результаты провалидированы при помощи технической спецификации PV модуля от производителя и относительная погрешность не превышает 1,65%. Также панель показывает отличные результаты на настоящих климатических данных (солнечное излучение, температура).

### **Выводы.**

Уже сейчас существует множество отдельных способов для решения проблемы мониторинга отдельных параметров и/или факторов солнечной PV панели, но отсутствуют комплексные решения по оценке всех факторов целиком и влиянию данных факторов на работу панели как в текущий момент, так и в ближайшем будущем при сохранении условий эксплуатации.

Различные математические модели позволяют оценить отдельные специфические параметры и связанные с ними риски, к примеру затенение - и также существуют решения для таких случаев, например методы реконфигурации панелей для минимизации ущерба от затенения и оптимизации производительности всех панелей на данной станции в целом. Однако по-прежнему актуальной является задача разработки методов и алгоритмов автоматизации поиска неисправностей солнечной электростанции на основе данных телеметрии или иным способом.

### **Список литературы**

- [1] Wang, H., Li, H., Tang, C. et al. Modeling, metrics, and optimal design for solar energy-powered base station system / J Wireless Com Network (2015) 2015: 39.
- [2] Kasaeian, A.B., Akhlaghi, M.M., Golzari, S. et al. Modeling and optimization of an air-cooled photovoltaic thermal (PV/T) system using genetic algorithms / Appl. Sol. Energy (2013) 49: 215.
- [3] Nguyen, X.H. & Nguyen, M.P. Mathematical modeling of photovoltaic cell/module/arrays with tags in Matlab/Simulink / Environ Syst Res (2015) 4: 24.
- [4] Hassan Ali, M., Rabhi, A., Haddad, . et al. Real-Time Determination of Solar Cell Parameters / Journal of Elec Materi (2017) 46: 6535.
- [5] Groenewolt, A., Bakker, J., Hofer, J. et al. Methods for modelling and analysis of bendable photovoltaic modules on irregularly curved surfaces / Int J Energy Environ Eng (2016) 7: 261.
- [6] Brahmi, H. & Dhifaoui, Dynamic characteristics and improved MPPT control of PV generator / R. Front. Energy (2013) 7: 342.
- [7] Kayri, I. & Gencoglu, M.T. Predicting power production from a photovoltaic panel through artificial neural networks using atmospheric indicators / Neural Comput & Applic (2017).
- [8] BHARADWAJ, P. & JOHN, V. Linearised model for PV panel power output variation with changes in ambient conditions / Sādhanā (2017) 42: 2183.
- [9] Abido, M.A. & Khalid, M.S. Seven-parameter PV model estimation using Differential Evolution / Electr Eng (2018) 100: 971.
- [10] Sarkar, M.N.I. Effect of various model parameters on solar photovoltaic cell simulation: a SPICE analysis / Renewables (2016) 3: 13.
- [11] Meflah, A., Rahmoun, K., Mahrane, A. et al. Outdoor performance modeling of three different silicon photovoltaic module technologies / Int J Energy Environ Eng (2017) 8: 143.
- [12] Altermatt, P.P. Models for numerical device simulations of crystalline silicon solar cells—a review / J Comput Electron (2011) 10: 314.

- [13] Abdelhamid, H., Edris, A., Helmy, A. et al. Fast and accurate PV model for SPICE simulation / *J Comput Electron* (2019) 18: 260.
- [14] Chaichan M.T., Kazem H.A. (2018) Environmental Conditions and Its Effect on PV Performance. In: *Generating Electricity Using Photovoltaic Solar Plants in Iraq*. Springer, Cham
- [15] Harrag, A. & Messalti, S. Adaptive GA-based reconfiguration of photovoltaic array combating partial shading conditions / *Neural Comput & Applic* (2018) 30: 1145.
- [16] Fialho L., Melício R., Mendes V.M.F., Figueiredo J., Collares-Pereira M. (2014) Amorphous Solar Modules Simulation and Experimental Results: Effect of Shading. In: Camarinha-Matos L.M., Barrento N.S., Mendonça R. (eds) *Technological Innovation for Collective Awareness Systems. DoCEIS 2014. IFIP Advances in Information and Communication Technology*, vol 423. Springer, Berlin, Heidelberg
- [17] Assouline D., Mohajeri N., Scartezzini JL. (2018) Estimation of Large-Scale Solar Rooftop PV Potential for Smart Grid Integration: A Methodological Review. In: Amini M., Borojeni K., Iyengar S., Pardalos P., Blaabjerg F., Madni A. (eds) *Sustainable Interdependent Networks. Studies in Systems, Decision and Control*, vol 145. Springer, Cham
- [18] Fezzani A., Mahammed I.H., Bazi S. (2017) A Mathematical Model to Determine the Shading Effects in the I-V Characteristic of a Photovoltaic Module. In: Chadli M., Bououden S., Zelinka I. (eds) *Recent Advances in Electrical Engineering and Control Applications. ICEECA 2016. Lecture Notes in Electrical Engineering*, vol 411. Springer, Cham
- [19] Vijayalekshmy, S., Rama Iyer, S. & Beevi, B. J. Comparative Analysis on the Performance of a Short String of Series-Connected and Parallel-Connected Photovoltaic Array Under Partial Shading / *Inst. Eng. India Ser. B* (2015) 96: 217.
- [20] LYDEN, S. & HAQUE, M.E. Modelling, parameter estimation and assessment of partial shading conditions of photovoltaic modules / *J. Mod. Power Syst. Clean Energy* (2019) 7: 55.
- [21] Exel L., Felgner F., Frey G. (2015) Efficient Models of Partially Shaded PV Modules for Energy System Design. In: Oral A., Bahsi Oral Z., Ozer M. (eds) *2nd International Congress on Energy Efficiency and Energy Related Materials (ENEFM2014)*. Springer Proceedings in Energy. Springer, Cham
- [22] Kahoul, N., Houabes, M. & Neçaibia, A. A comprehensive simulator for assessing the reliability of a photovoltaic panel peak power tracking system / *Front. Energy* (2015) 9: 170.
- [23] Kahoul, N., Houabes, M. & Neçaibia, A. A comprehensive simulator for assessing the reliability of a photovoltaic panel peak power tracking system / *Front. Energy* (2015) 9: 170.
- [24] Davarifar M., Rabhi A., EL Hajjaji A., Bosche J., Pierre X. (2013) Improved Real Time Amorphous PV Model for Fault Diagnostic Usage. In: Hakansson A., Höjer M., Howlett R., Jain L. (eds) *Sustainability in Energy and Buildings. Smart Innovation, Systems and Technologies*, vol 22. Springer, Berlin, Heidelberg
- [25] Tina G., Cosentino F., Ventura C. (2016) Monitoring and Diagnostics of Photovoltaic Power Plants. In: Sayigh A. (eds) *Renewable Energy in the Service of Mankind Vol II*. Springer, Cham
- [26] Arias-Rosales, A. & Mejía-Gutiérrez, R. Modelling and simulation of direct solar radiation for cost-effectiveness analysis of V-Trough photovoltaic devices / *Int J Interact Des Manuf* (2016) 10: 257
- [27] Vinod, Raj Kumar, S.K. Singh. Solar photovoltaic modeling and simulation: As a renewable energy solution / *Energy Reports*, Volume 4, 2018, Pages 701-712, ISSN 2352-4847.

## **ALGORITHMS FOR TROUBLESHOOTING A SOLAR POWER PLANT BASED ON BIG DATA OF ITS MONITORING**

***K.S. Dick***

*Postgraduate student of the  
Belarusian State University of  
Informatics and Radioelectronics*

***I.I. Piletsky***

*Associate Professor of the Department of  
Informatics of the Belarusian State University of  
Informatics and Radioelectronics, Ph.D*

*Department of Informatics  
Belarusian State University of Informatics and Radioelectronics*

**Abstract.** The analysis of algorithms for troubleshooting solar power plants based on big data collected as a result of monitoring using cloud resources is presented. Various approaches to modeling solar panels and solar cells are analyzed, feedback control systems and the influence of changes in environmental parameters on the performance of solar power plants are investigated. Special attention is paid to comparing different models of solar cells to determine the influence of input parameters on their efficiency.

**Keywords:** anomaly detection, solar PV panel, deep learning, unsupervised learning, mathematical models.

УДК 332.14:338.001.36

## ЦИФРОВОЙ УНИВЕРСИТЕТ 4.0 КАК ОБРАЗОВАТЕЛЬНАЯ СОЦИОТЕХНИЧЕСКАЯ ЭКОСИСТЕМА ДЛЯ ИНДУСТРИИ 4.0



**А.Г. Давыдовский**  
доцент кафедры программного обеспечения информационных технологий факультета компьютерных систем и сетей БГУИР, кандидат биологических наук, доцент  
agd2011@list.ru



**Л.Н. Воронцовая**  
проректор по научно-методической работе ГУО МГИРО, кандидат педагогических наук, доцент  
voroneckaya@minsk.edu.by



**А.В. Пицова**  
доцент кафедры социальной педагогики факультета социально-педагогических технологий БГПУ имени Максима Танка, кандидат педагогических наук, доцент  
anita\_17@list.ru

### **А.Г. Давыдовский**

Область научных интересов связана с проблемами социальной информатики, системного анализа, управления и обработки информации, математического моделирования биосоциальных и социотехнических систем, методологии превентивного управления рисками в биосоциальных, образовательных и социотехнических системах. Автор монографии, учебных программ и пособий для студентов и магистрантов.

### **Л.Н. Воронцовая**

Область научных интересов связана с перспективами развития и реализации модели университета 4.0 и ее версий, а также проблемами воспитания, народной педагогики, формирование национального самосознания, патриотического воспитания. Автор монографии, учебных программ и пособий для студентов и магистрантов.

### **А.В. Пицова**

Область научных интересов связана с проблемами социальной информатики, медиабезопасности субъектов образовательного процесса, а также социокультурного анализа образовательных систем. Научный руководитель инновационной студенческой научно-исследовательской лаборатории «Стратегии и модели социализации молодежи XXI века» ФСПТ БГПУ. Автор учебных программ и пособий для студентов и магистрантов.

**Аннотация.** Разработаны основы концепции модели цифрового университета 4.0 как образовательная социотехническая экосистема в условиях четвертой промышленной революции. Выделены и охарактеризованы пять иерархических уровней структурно-функциональной организации цифрового университета 4.0. Рассмотрены особенности и перспективы реализации цифрового университета 4.0.

**Ключевые слова:** цифровой университет 4.0, структурно-функциональная организация, социотехническая экосистема, Индустрия 4.0.

### **Введение.**

В историческом отношении университет как социокультурный феномен прошел, по меньшей мере, три этапа эволюционного развития, каждый из которых может быть охарактеризован особой моделью выполняемой социокультурной миссии и взаимоотношений с внешней социокультурной, производственно-технологической и социально-экономической средой. Эти этапы, последовательно сменявшие друг друга, могут быть сравнительно условно описаны моделями университета 1.0, 2.0 и 3.0. При этом информационная метафора в цифровом обозначении позволяет охарактеризовать лишь число основных социокультурных миссий, выполняемых университетом. В частности, если университет 1.0 – это только лишь образовательный институт, то университет 2.0 обеспечивает поддержание уже двух миссий – как

образования, так и проведение научных исследований, тогда как в университете 3.0 эти две миссии интегрируются с инновационной деятельностью с последующей коммерциализацией новых знаний и технологий, вновь разработанных на их основе. Причем университет 3.0 становится драйвером социально-экономической модернизации ряда многих секторов национальной экономики, а также трансформации общества в целом на пути перехода к экономике знаний [1, 2].

В первой четверти XXI века переход к четвертой промышленной революции (Индустрии 4.0) и экономике знаний обусловил необходимость радикальной трансформации университетского образования. Все большее значение приобретают функции университета, связанные с практико-ориентированными исследованиями, опытно-конструкторскими разработками, инновационной и коммерческой деятельностью, а также многовекторным развитием человеческого потенциала всех участников образовательного процесса. Университет все чаще рассматривается как центр инновационного развития территорий. В этой связи особую актуальность приобретает разработка и реализация новых развития университетского образования в направлении университета 4.0 [3, 4].

При этом университет 4.0 является инновационным научно-образовательным комплексом, включающим многочисленные компоненты и направления деятельности, которые обладают образовательными информационными, технологическими, материальными ресурсами, а также человеческим потенциалом для решения образовательных, научно-исследовательских, опытно-конструкторских, производственно-технологических, технологических, инновационно-коммерческих задач, принципиально нереализуемых в традиционных моделях социально-экономических и производственных отношений. Это, прежде всего, относится к задачам развития, инсталляции и использования конвергентных информационных, нано-, био- и когнитивных технологий. Такие задачи являются вызовами, характерными для нынешнего этапа интеллектуально-технологического развития цивилизации, когда изменяется структура передового производства, а центр тяжести интеллектуальных усилий смещается в область проектирования, инновационных разработок и их экспериментального освоения. Очевидно, рост производительных сил, обусловленных развитием четвертой промышленной революции, будет неизбежно сопровождаться совершенствованием производственных отношений, что потребует смены образовательной парадигмы в следующей модели университетского образования – университета 4.0 как кластерной образовательной платформы, решающей не только долгосрочные образовательные, научно-исследовательские, проектно-инновационные и инновационно-коммерческие программы, а также обеспечивающей развитие человеческого потенциала для вызовов будущего [5, 6].

Цифровая трансформация экономики, переход от глобализации к глокализации, сочетающей в себе глобальные и региональные экономические, социальные, культурные, политические процессы и тенденции на основе новых возможностей, предоставляемых цифровыми технологическими платформами, позволяет перейти к индивидуальному освоению образовательных программ, когда обучающиеся могут осваивать образовательные онлайн-программы в удобном для себя темпе, в любое время вне зависимости от часового пояса и географической удаленности от вуза, существенно снизить затраты на печатные материалы в связи с тем, что учебно-методическое обеспечение образовательного процесса полностью существует в цифровом формате, широкому использованию технологий дистанционного обучения, которые позволяют не увеличивать затраты на масштабирование образовательных услуг независимо от количества обучающихся.

С другой стороны, переход к модели цифрового университета или «университета будущего» предъявляет к преподавателям, студентам и руководству образовательной организации ряд принципиально новых компетенций. Согласно А.В. Курдюмову, «системой ключевых компетенций образовательного учреждения является перечень факторов, способствующих повышению конкурентоспособности вуза» [7].

Первоочередными мероприятиями, направленными на обретение университетом 4.0 ключевых компетенций, повышающих его конкурентоспособность, являются [8, 9]:

- необходимость перманентного повышения квалификации всего профессорско-преподавательского состава, в первую очередь в области информационно-коммуникационных технологий (ИКТ);
- разработка системы стимулирования преподавателей, активно внедряющих цифровые методы в образовании и активно развивающих свои собственные навыки;
- активная поддержка преподавателей, в наименьшей степени обладающих необходимыми базовыми знаниями в области ИКТ;
- необходимость кардинальной перестройки инфраструктуры и бизнес-процессов университета 4.0;
- внедрение технологий управления индивидуальной образовательной траектории как для обучающихся, так и для преподавателей;
- переход на цифровые методы образования MOOCs (Massive Open Online Courses), удаленные методы рубежного контроля.

В этой связи целью исследования является разработка и анализ основных положений концепции цифрового университета 4.0 как социотехнической экосистемы в условиях четвертой промышленной революции.

#### **Основы концепции цифрового университета 4.0 как социотехнической экосистемы.**

Необходимо отметить, что одной из наиболее перспективных моделей развития высшей школы в условиях цифровой трансформации в условиях четвертой промышленной революции является разработка и реализация модели цифрового университета 4.0, включающей развитие комплекса ИКТ-платформ и аналитических приложений, направленных на приращение когнитивного потенциала всех участников образовательного процесса [10]. Причем цифровой университет 4.0 является социотехнической системой, включающей, как минимум, три основных компонента: антропосоциальный, технологический и внешнесредовой, представленный комплексом социокультурных, социально-экономических, историко-политических и других факторов. Очевидно, целевой установкой цифрового университета 4.0 может быть поиск и развитие человеческого потенциала и smart-моделирование сценариев профессиональной, творческой и личной биографии человека на основе технологий искусственного интеллекта, информатики, графического анализа и прогнозирования, а также технологий ноосферного развития. Цифровые приложения (скрипты) полностью адаптируются под потребности человека и окончательно вытеснят классические образовательные программы и линейный способ передачи информации. При этом студенты смогут обучаться в любом месте и в любое время, независимо от возраста, социальной и профессиональной принадлежности, культурных, этических и идеологических предпочтений. В этих условиях значительно редуцируется традиционное дисциплинарное ядро современного образования, на смену которому приходит проблемно-ориентированное трансдисциплинарное образование на основе проектных, личностно ориентированных и конвергентных образовательных технологий. При этом ключевыми категориями цифрового университета 4.0 как образовательной социотехнической экосистемы являются развитие творческого потенциала человека, выступающего в роли творца и созидателя, а также инновации в качестве ключевого регулятора межинституциональных отношений [1, 5, 6, 10].

Цифровизация обеспечивает возможности для обмена накопленным опытом и знаниями, что позволяет людям узнать больше и принимать более обоснованные решения в своей повседневной жизни. В актуальной ситуации прослеживаются очевидные тенденции развития университетов и высшего образования, среди которых изменение статуса университетов за счет усиления в научной деятельности рисков и экспериментирования, переход от конкуренции к партнерскому взаимодействию, обращение к большим базам данных (Big Data), переход к разноформатным открытым образовательным ресурсам (Open Online Resources), сочетание

новых и традиционных форматов подготовки, перепланировка учебных помещений в формате открытых пространств (Open Space). Распространение получают такие технологии, как обучение с использованием мобильных устройств (Bring Your Own Devices), переход к «перевернутым классам» (Flipped Classroom), создание «конструктора пространств» (Makerspaces) – высокотехнологичных площадок с использованием 3D принтеров, создание мобильных и адаптивных технологий обучающихся, развитие адаптивного обучения за счет внедрения цифровых платформ и распространение интернета вещей. Дополнительные направления применения цифровых технологий в образовании – развитие онлайн-цифровых библиотек и цифровых кампусов университетов, которые уже внедрены многими университетами в Америке, Европе, России. Среди интересных цифровых инноваций следует отметить быструю адаптацию онлайн-обучения, которое выражается в виде развития смешанных форм обучения (blended learning) и в активном развитии онлайн-курсов MOOC (Massive on-line open course). Наступающая цифровая эпоха потребует от всех участников – университетов, локальных сообществ, государств, международных организаций, транснациональных компаний – интеграции действий. В свою очередь, в новых технологических реалиях существенно возрастает роль непрерывного дополнительного образования взрослых.

Сфера образования, наряду с рядом других отраслей (таких, как здравоохранение и телекоммуникации) подвержена существенным изменениям из-за все более активного распространения цифровых технологий в условиях развития тенденций четвертой промышленной революции. Как обычно, тренды в области внедрения цифровых технологий в образовательную и научно-исследовательскую деятельность задают коммерческие организации: частные университеты, бизнес-школы, корпоративные университеты, а также государственные технические университеты ведущих стран. [9, 11].

Цифровизация системы университетского образования является формой реализации проектно-конструкторского подхода в сфере управления сценариями будущего развития сферы образования. Перед университетами, стремящимися сохранить свои позиции на глобальном рынке образования, стоит задача вхождения в международное научно-образовательное пространство. Цифровые платформы, институционально замещающие прежние каналы хранения, обработки и получения информации, «умные технологии», цифровые приложения и образовательные ресурсы MOOC (Massive Open Online Courses), SPOOC (Self-Paced Online Courses) в условиях нерационального использования являются деструктивными факторами мирового образовательного пространства, постепенно изменяющими способы получения образования и разрушающими привычные формы его организации [1, 9].

Таким образом, изменения идентичности университетского образования приобретают необратимый характер, неизбежность реструктуризации формата его концептуальной модели становится очевидной.

На основе обобщения и анализа международного опыта использования организации и реализации различных форм образовательной деятельности в сфере дополнительного образования была разработана концептуальная модель сетевой образовательной платформы цифрового университета 4.0, направленная на многомерное и многовекторное развитие человеческого потенциала обучающихся.

При этом аналитические приложения и образовательные ресурсы, адаптированные под потребности и возможности человека, фундаментально меняют способы получения образования, следовательно, его архитектуру, принципы, цели и сущностные характеристики. Возникает феномен «образования по требованию» (Learning on Demand), когда образовательный продукт конструируется под субъекта [9, 11].

Принимая во внимание вышеизложенное, предложены основы концепции социотехнической экосистемы цифрового университета 4.0 как оптимально соответствующего трендам и вызовам Индустрии 4.0.

Согласно данной концепции, в условиях Индустрии 4.0 глобальной цифровизации



образования, включая ее отдельные негативные последствия, социотехническая система цифрового университета 4.0 включает интеграцию ИКТ-сервисов, цифровых образовательных ресурсов, комплекса факторов человеческого потенциала, в частности, академические, социально-личностные, профессионально ориентированные и цифровые компетенции. В рамках предложенной модели цифрового университета 4.0 цифровые приложения (скрипты) полностью адаптируются под потребности человека и окончательно вытеснят классические образовательные программы и линейный способ передачи информации. Студенты смогут обучаться в любом месте и в любое время. В этих условиях можно предположить переход от дисциплинарного образования к трансдисциплинарному образованию, ориентированному на освоение тематических блоков с последующим решением соответствующих практико-ориентированных комплексных проблем. Университет прекратит традиционное существование в физическом пространстве, расширяя свое присутствие и влияние в виртуальных пространствах цифрового университета Интернет посредством облачных технологий, включая возможности формата сетевого партнерства с распределенным управлением. Транснациональные компании и общественные организации смогут включаться в социальный заказ для подготовки специалистов с заданным набором компетенций.

Университет 4.0 может рассматриваться как универсальный образовательный сервис в условиях Индустрии 4.0 и в дальнейшем может позиционироваться как корпоративный субъект экономики знаний, осуществляющий образовательную, научно-исследовательскую, инновационную и коммерческую деятельность в условиях «информационного взрыва» с экспоненциальным ростом объема, семантической сложности и структурного разнообразия информационных потоков [1, 5].

Изучение проблем управления техническим университетом 4.0 как социотехнической экосистемой в условиях Индустрии 4.0 и перехода к VI технологическому укладу требует выделения важнейших характеристик для его системного анализа и моделирования, включая:

- полифункциональность образовательной деятельности (МОД) университета 4.0, включая дифференциацию и диверсификацию образовательной подготовки студентов;
- консалтинговая деятельность и трансфер знаний (КДТЗ);
- коммерциализация инновационных технологий (КИД);
- многоуровневая система высшего образования (МСВО);
- высокое качество образования (ВКО) бакалавров и магистров, а также научных кадров высшей квалификации (кандидатов и докторов наук), может быть индивидуально оптимизировано на основе индивидуального образовательного маршрута;
- индексные показатели относительного баланса контингентов обучающихся (БКО) на различных уровнях университета, включая соотношение магистрантов к бакалаврам, аспирантов к магистрантам, докторантов к аспирантам, причем важнейшее значение для устойчивого развития цифрового университета 4.0 как образовательной социотехнической экосистемы имеет оптимизация этих соотношений;
- направления образовательной деятельности (НОД), соответствующих инновационным научным, технологическим и социально-гуманитарным направлениям для опережающего развития наиболее конкурентоспособных отраслей национальной экономики с учетом результатов системного прогнозирования мировых и национальных трендов социально-экономического и научно-технологического развития;
- интеграция потоков административно-управленческой информации (АУИ) на основе облачных технологий и Big Data [11].

Структурно-функциональную организацию цифрового университета 4.0 (ЦУ 4.0) как образовательной социотехнической экосистемы можно представить кортежной моделью (1):

$$\text{ЦУ 4.0} = \langle \text{МОД, КДТЗ, КИД, МСВО, ВКО, БКО, НОД, АУИ} \rangle. \quad (1)$$

Важной проблемой управления цифровым университетом 4.0 как социотехнической системой является переход к принципиально новой форме интеграции всех его иерархических уровней и функций, включая применение электронных дидактических технологий, обеспечивающих повышение качества образовательного процесса, академической успеваемости обучающихся на основе оптимизации индивидуальных образовательных маршрутов обучающихся.

#### **Структурно-функциональная организация социотехнической экосистемы цифрового университета 4.0.**

При формировании университета 4.0 целесообразно придерживаться пятивекторной модели цифровой трансформации, согласно которой радикальные изменения должны произойти с пятью сферами жизнедеятельности университета, включая [12]:

- системы управления на основе обработки потоков «больших данных» для создания интегральной инновационной экосистемы университета;
- цифровизация учебно-методического сопровождения образовательного процесса, образовательных технологий и коммуникаций между его субъектами;
- формирование и сопровождение индивидуальных образовательных траекторий;
- развитие цифровой технологической и «облачной» инфраструктуры университета;
- формирование, развитие и селекция человеческого потенциала на основе перманентной поддержки цифровых компетенций всех акторов образовательного процесса университета 4.0 (студентов, преподавателей, администрации);

Переход к цифровой инфраструктуре социотехнической экосистемы цифрового университета 4.0 описывает комплекс мероприятий по глобальной перестройке всех бизнес-процессов и преобразование структурных подразделений образовательной организации на основе использования ИКТ и единой информационно-образовательной системы цифрового университета 4.0 [14].

Принимая во внимание высокую сложность цифровой трансформации цифрового университета 4.0, на начальных этапах целесообразно использовать принцип *raref-free* (безбумажный документооборот), модуль мониторинга бизнес-процессов университета 4.0, системе одного окна и мобильной организационной и ИКТ-инфраструктуры университета как научно-образовательной бизнес-корпорации [15, 16].

Внедрение концепции безбумажной обработки данных является основой системы электронного документооборота для автоматизации системы прохождения и согласования документов внутри образовательной организации университета 4.0. При этом минимальная доля рабочих документов, проходящих через систему автоматизации документооборота, должно составлять не менее 80 %.

В свою очередь, модуль мониторинга бизнес-процессов в режиме реального времени позволяет с помощью технологий обработки и анализа «больших данных» осуществлять сбор и анализ деятельности кафедр и институтов (факультетов) университета 4.0, собирать сведения о его научно-исследовательской деятельности, мониторить системы безопасности, включая камеры видеонаблюдения, устройства контроля доступа на территорию, пожарной безопасности, функционирование локальной сети и коммунальной инфраструктуры.

Система одного окна и мобильной инфраструктуры цифрового университета 4.0 обеспечивает преодоление внутренних бюрократических барьеров для коммуникаций между студентами, преподавателями, сотрудниками администрации.

Кроме того, в рамках концепции цифрового университета 4.0 как социотехнической экосистемы целесообразно выделить мероприятия, относящиеся к следующим пяти уровням [4, 6, 11, 13, 18].

Первый уровень (основной) – включает слушателей и представлен непосредственно самими слушателями, корпусом профессорско-преподавательского состава, научно-инженерными работниками, отраслевыми, академическими, научно-производственными и

научно-исследовательскими организациями-партнерами, которые являются внутренними и внешними стейкхолдерами цифрового университета 4.0.

Второй уровень (информационно-ресурсный) включает в себя информационно-технологические сервисы, которые являются обязательным элементом цифрового университета 4.0. Этот уровень представлен базовыми информационными сервисами. Их задача – создание единого информационного пространства для цифрового взаимодействия внутри цифрового университета 4.0 с использованием гибких smart-инструментов, включая видеозкраны для проведения лекций и семинаров, беспроводную связь на всей территории университета, защищенные облачные хранилища для хранения и обмена данными, другие технологии, соответствующие тенденциям Индустрии 4.0. Цифровая библиотека обеспечивает доступ слушателя или преподавателя к электронным ресурсам учебно-методической и научной литературы с любых устройств, независимо от места нахождения и времени суток. Цифровизация наукометрии состоит в мониторинге, накоплении и анализе наукометрической информации с использованием современных методов хранения и обработки больших массивов данных. Многие современные университеты объединяют традиционные и цифровые библиотеки с точки зрения опыта конечного пользователя. Подобная конвергенция традиционных и новых технологий обеспечивает более высокий уровень комфорта для студентов и преподавателей и позитивно влияет на имидж цифрового университета 4.0.

Третий уровень (процессный) является наиболее ресурсоемким, поскольку включает цифровой маркетинг, управление исследовательскими проектами, управление закупками, взаимодействие со слушателями. Вместе с тем, это позволяет информационно-образовательной платформе цифрового университета 4.0 получить наибольшую добавленную стоимость.

В частности, цифровой маркетинг является новой для классических, инновационных и технических университетов областью деятельности, направленной на решение таких задач, как:

- организация взаимодействия со слушателями с использованием всего современного спектра цифровых каналов коммуникации;
- мониторинг изменений в восприятии бренда университета на целевых рынках на основе результатов исследований и мониторинга онлайн-социальных сетей и мессенджеров;
- проведение превентивных и реактивных мероприятий для формирования положительного имиджа информационно-образовательной платформы цифровой университет 4.0;
- стимулирование создания новых цифровых сообществ и инноваций на всех этапах образовательного цикла, а также коммуникации содержания образовательных программ и особенностей образовательной деятельности для слушателей;
- разработка персонализированных маркетинговых материалов для целевых аудиторий на основе анализа данных из разных источников.

Взаимодействие со слушателями включает следующие задачи:

- использование цифровых технологий для взаимодействия со слушателями повышения квалификации и переподготовки;
- разработка цифровых моделей индивидуальных образовательных маршрутов слушателей повышения квалификации и переподготовки;
- использование различных цифровых каналов коммуникации для предоставления слушателям дистанционного доступа к образовательным ресурсам;
- использование методов аналитики Big Data и Data Mining для определения наиболее успешных и перспективных слушателей;
- организация и управление самоорганизующимися «онлайн-сетевыми сообществами слушателей».

Четвертый уровень (образовательный) включает информационные и материально-технические инструменты, направленные на освоение представлений и знаний по направлениям

четвертой промышленной революции, которые с высокой степенью вероятности получат широкое распространение в университетской среде во второй четверти XXI века. К таким технологиям, например, относятся smart-технологии, искусственный интеллект, беспилотный транспорт, конвергентные нано-, био-, информационные и когнитивные технологии, аддитивные технологии, технологии виртуальной и дополненной реальности, блокчейн, когнитивные образовательные технологии, квантовые вычисления, индустриальное освоение космоса, социальная инженерия и социоинженерное проектирование, технологии кибербезопасности, превентивное управление рисками природных и техногенных катастроф, организационные технологии, технологии сетецентрического управления социотехническими системами, системные smart-технологии, технологии «Интернета вещей».

Пятый уровень (поддерживающий) включает мероприятия, использование которых предусмотрено на всех этапах реализации проекта цифрового университета 4.0 как социотехнической экосистемы:

–отслеживание технологических новинок и консультирование по вариантам их возможного использования для достижения поставленных перед университетом целей;

–совершенствование политики и процедур, направленных на стимулирование использования инновационных цифровых технологий среди административного персонала вуза, обучающихся и научно-педагогических работников;

–предоставление максимального открытого и удобного доступа к информационным ресурсам и системам с целью обеспечить возможность использования данных посредством новых технологий;

–оптимизация использования облачных решений для стимулирования инноваций и быстрой оборачиваемости нового цифрового функционала, продуктов и систем;

–реализация процессов повышения квалификации и переподготовки наряду с процессами научной деятельности с максимальным использованием потенциала цифровых технологий;

–разработка индивидуально оптимизированных программ непрерывного повышения квалификации для обеспечения постоянного развития цифровых компетенций профессорско-преподавательского состава, технического персонала и администраторов цифрового университета 4.0 на основе ротационного принципа.

### **Заключение.**

Реализация концепции цифрового университета 4.0 как образовательной социотехнической экосистемы будет стимулировать коэволюционное развитие индивидуальных возможностей как отдельной человеческой индивидуальности, так и коллективного человечества на основе техногуманитарного синтеза культурно-информационных ресурсов человеческой цивилизации и высоких технологий. Такая трансформация заключается не только и столько во внедрении цифровых решений, сколько в целом является существенным культурным и организационным изменением в самой образовательной парадигме о роли университета как социокультурного феномена в потоке цивилизационных событий. Очевидно, трансформация высшей школы на платформе цифрового университета 4.0 представляется необходимым инструментом перехода к ноосферному миру после мира многополярного, развивающегося на основе комплекса идей глокализации и конкуренции между отдельными техноэкономическими региональными блоками.

Таким образом, реализация модели социотехнической экосистемы цифрового университета 4.0 предоставляет ряд преимуществ на рынке образовательных услуг вследствие опережающего характера и оптимизации управления образовательной деятельностью, повышения качества образования, интенсификации организационно-управленческих изменений, разработки и внедрения принципиально новых учебных дисциплин и практик, а также индивидуальных образовательных маршрутов студентов, магистрантов и аспирантов для

оптимизации их индивидуальной перманентной образовательной деятельности в условиях четвертой промышленной революции.

Переход к цифровому университету 4.0 предполагает внедрение более гибких и бесшовных процессов, изменение корпоративной культуры, оптимизацию процессов. Предлагаемая концепция цифрового университета 4.0 радикально отличается от идей и сценариев трансгуманистического и постгуманистического развития человечества, рассматривающих человека разумного (*Homo sapiens*) как тупиковую ветвь логической эволюции и предусматривающих сначала переход от *Homo sapiens* к человеку кибернетическому (*Populus cyber*), вплоть до полной замены человека самовоспроизводящимся и самодостаточным искусственным разумом [9, 11, 13, 18]. Возможно, предложенная концепция цифрового университета 4.0 позволит сделать один из необходимых шагов в направлении избегания реализации трансгуманистических и постгуманистических сценариев трансформации высшего образования, которые по сути, являются сценариями не только отмены образования, но и глубинной модификации самой природы человека, которая будет иметь, несомненно, трагические последствия. Реализация концепции цифрового университета 4.0 позволит создать первоначальные предпосылки для перехода человека разумного (*Homo sapiens*) к человеку ноосферному (*Homo de noosphere*) на основе технугуманитарной коэволюции логических и психофизиологических возможностей человека разумного.

#### Список литературы

- [1] Ефимов, В. С. Фазовые трансформации и будущее университетов: философско-методологический анализ / В.С. Ефимов, А.В. Лаптева // Университетское управление: практика и анализ. – 2016. – №6. – С. – 146–158.
- [2] Карпов, А. О. Современный университет как драйвер экономического роста: модели и миссии / А.О. Карпов. – [Электронный ресурс]. – Режим доступа : <http://institutiones.com/general/2975-sovremennyi-universitet.html>. – Дата доступа : 19.01.2019.
- [3] Шваб, К. Четвертая промышленная революция / К. Шваб. – М. : Эксмо, 2016. – 278 с.
- [4] Curley, M. University ecosystems design creative spaces for idea generation and start-up experimentation / M. Curley, P. Formica // In: M. Curley, P. Formica (eds.). The experimental nature of new venture creation : Capitalizing on open innovation 2.0. N.Y.: Springer, 2013. – 13–23.
- [5] Цифровой университет: применение цифровых технологий в современных образовательных учреждениях. – [Электронный университет]. – Режим доступа: [https://www.itweek.ru/idea/article/detail\\_print.php?ID=192831&print=Y](https://www.itweek.ru/idea/article/detail_print.php?ID=192831&print=Y). – Дата доступа: 12.07.2019.
- [6] Модель для сборки университетов 4.0, или Для кого не наступит будущее. Что делать вузам, чтобы выжить в меняющемся мире. – [Электронный ресурс]. – Режим доступа : <https://indicator.ru/article/2017/09/22/konferenciya-5-100-university>. – Дата доступа : 08.05.2018.
- [7] Гольшкова И.Н. Анализ ключевых составляющих модели «Цифровой университет»//E-Management. 2020. № 3. С. 53-61.
- [8] Безуглова, И.Г. (2018). Ключевые компетенции, как основа устойчивого конкурентного преимущества образовательной организации / И.Г. Безуглова [Электронный ресурс]. – Режим доступа: <https://nsportal.ru/shkola/administrirovanie-shkoly/library/2016/02/05/klyuchevye-kompetentsii-kak-osnova-ustoychivogo>. – Дата доступа: 12.04.2023.
- [9] Давыдовский, А.Г. Проблемы управления социально-экономической экосистемой инновационного научно-образовательного комплекса «технический университет 4.0» / А.Г. Давыдовский // Актуальные теоретические и прикладные вопросы управления социально-экономическими системами: Материалы Международной научно-практической конференции. – Том. 3. Москва, 20 декабря 2019. – М.: Институт развития дополнительного профессионального образования, 2019. – С. 18–28.
- [10] Неборский, Е.В. Реконструирование модели университета: переход к формату 4.0 / Е.В. Неборский // Мир науки. – 2017. – Том 5, N 4. – [Электронный ресурс]. – Режим доступа : <http://mir-nauki.com/PDF/26PDMN417.pdf>. – Дата доступа : 17.05.2019.
- [11] Давыдовский, А.Г. Системный анализ инновационного и научно-образовательного комплекса университета 4.0 / А.Г. Давыдовский, Н.В. Лапицкая, М.И. Морозова, А.В. Пищова // В сборнике материалов V Международной научно-практической конференции «BIG DATA и анализ высокого уровня». В двух частях. – 2019. – Часть 2. – С. 325–336.
- [12] Dietz J., Hoogervorst J. et al. The Discipline of Enterprise Engineering// International Journal of Organisational Design and Engineering. – 2013. – Vol.3, N1.– P.86–114.
- [13] Давыдовский, А.Г. Социотехнический анализ образовательной системы университета 4.0 / А.Г.

Давыдовский // Системное моделирование социально-экономических процессов: труды 42-ой Международной научной школы-семинара, г. Ростов-на-Дону, 1 – 6 октября 2019 г. / под ред. д-ра экон. наук В.Г. Гребенникова, д-ра экон. наук И.Н. Щепиной. – Воронеж: Изд-во «Истоки», 2019. – (648 с.) – С. 617–621.

[14] Wissema, J. Towards the third generation University. Managing the University in transition / J. Wissema (1st ed.). – Edward Elgar Publishing Limited, 2010. – 272 p.

[15] Ефимов, В.С. Университет 4.0: философско-методологический анализ / В.С. Ефимов, А.В. Лаптева // Университетское управление: практика и анализ. – 2017. – Т. 21, N 1. – С. 16–29.

[16] Пономаренко, Е.В. Новые модели развития университетов в мире в условиях цифровой революции: теоретические и практические подходы / Е.В. Пономаренко // Государственная служба. – 2017. – Т. 19, N6 (110). – С. 57–63.

[17] Флек, М.Б. Формирование модели инженера в условиях цифровой трансформации: взаимодействие предприятий и университетов / М.Б. Флек, Е.А. Утнич // Управление персоналом и интеллектуальными ресурсами в России. – 2018. – Т. 7, N4. – С. 45–52.

[18] Давыдовский, А.Г. Социальные последствия цифровой трансформации образования в условиях коронакризиса / А.Г. Давыдовский, Н.В. Лапицкая // Журнал Белорусского государственного университета. Социология. – 2022. – № 2. – С. 56-65.

## **DIGITAL UNIVERSITY 4.0 AS AN EDUCATIONAL SOCIOTECHNICAL ECOSYSTEM FOR INDUSTRY 4.0**

**A.G. Davidovsky**

*Associate Professor of the  
Department of Information  
Technology Software of the Faculty  
of Computer Systems and Networks  
of BSUIR, Candidate of Biological  
Sciences, Associate Professor*

**L.N. Voronetskaya**

*Vice-Rector for Scientific and  
methodological work of the State  
Educational Institution "Minsk City  
Institute of Education Development",  
Candidate of Pedagogical Sciences,  
Associate Professor*

**A.V. Pishchova**

*Associate Professor of the  
Department of Social Pedagogy of  
the Faculty of Social and  
Pedagogical Technologies of Maxim  
Tank BSPU,  
Candidate of Pedagogical Sciences,  
Associate Professor*

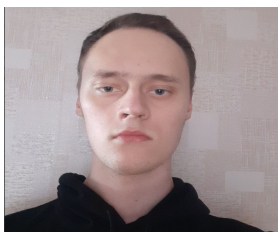
*Department of Information and Computer Systems Design  
Faculty of Computer Engineering Department of Information Technology Software  
of the Faculty of Computer Systems and Networks of  
Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: agd2011@list.ru*

**Abstract.** The fundamentals of the concept of the digital university 4.0 model as an educational sociotechnical ecosystem in the conditions of the Fourth Industrial Revolution have been developed. Five hierarchical levels of the structural and functional organization of the digital university 4.0 are identified and characterized. The features and prospects of the implementation of the digital university 4.0 are considered.

**Keywords:** digital university 4.0, structural and functional organization, sociotechnical ecosystem, Industry 4.0.

УДК 004.418:336.7

## BIG DATA В КРАУДФАНДИНГЕ КАК СРЕДСТВО ПРОГНОЗИРОВАНИЯ УСПЕХА КАМПАНИИ



**С.Н. Барсукевич**

Студент 4 курса факультета компьютерных систем и сетей специальности ИиТП БГУИР  
mx2002bsn0307@gmail.com



**С.Н. Нестеренков**

Декан факультета компьютерных систем и сетей БГУИР, кандидат технических наук, доцент  
s.nesterenkov@bsuir.by



**Д.В. Низовцов**

начальник отдела информационных технологий, ОИТ ассистент, Каф.ПОИТ  
d.nizovtsov@bsuir.by

### **С.Н. Барсукевич**

Студент 4 курса специальности “Информатика и Технологии Программирования” Белорусского государственного университета информатики и радиоэлектроники.

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцента кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации

### **Д.В. Низовцов**

Окончил БГУИР в 2016 году по специальности “Программное обеспечение информационных технологий”, магистрант первого года обучения по специальности “Электронные системы и технологии” БГУИР.

**Аннотация.** В данной работе представлен краткий обзор сферы краудфандинга, проанализированы важные аспекты успеха кампаний, а также определены пути использования больших данных для прогнозирования положительного результата.

**Ключевые слова:** краудфандинг, анализ данных, успех кампании, мониторинг.

### **Введение.**

Краудфандинг существует довольно давно. Этот процесс представляет собой сбор средств для различных проектов, целей или разработок, и это может включать огромные денежные суммы. Поэтому точные прогнозы для таких кампаний важны. Соответственно анализ больших данных является важным инструментом для прогнозирования успеха краудфандинговых кампаний [1].

Краудфандинг включает в себя три основных объекта: инициатор, который предлагает свои мысли общественности; сторонники идеи или проекта, которые на самом деле помогают инициатору, собирая деньги; и посредник, который объединяет обе стороны и выступает в качестве канала связи [2].

Одной из самых больших проблем в области краудфандинга является вероятность успеха проекта. Если инвесторам проект недостаточно нравится или они не в состоянии понять идею, они не проявят особого интереса к проекту. Недостаточная видимость и неудачная реклама (или ее полное отсутствие) значительно влияют на рост проекта: кампания никогда не будет замечена или не будет привлекательной для целевой аудитории. Таким образом, проект не получает должного старта, а реализация инициатора полностью приостанавливается из-за недостаточного

финансирования. Такие риски следует выявлять на ранней стадии проекта, чтобы их влияние на дальнейшую судьбу проекта не стало критичным.

#### **Основная часть.**

Определим какую ценность и влияние имеют большие данные для кампаний. Помимо всего прочего, большие данные могут помочь инициаторам кампаний получить ответы на наиболее важные вопросы, а также получить необходимую информацию по таким темам, как:

- Каковы в настоящее время наиболее популярные типы проектов?
- Каков статус подобных проектов и были ли они успешными?
- Если они были успешными, то какой метод краудфандинга они использовали?
- Каков был размер их базы сторонников?

Эти вопросы очень важны, и на них необходимо получить правильные ответы, прежде чем запускать какую-либо краудфандинговую кампанию. Большие данные могут произвести революцию в мире краудфандинга, предоставляя информацию о различных областях, делая каждую краудфандинговую кампанию успешной и полезной для общественности.

Big Data позволяет проанализировать множество факторов, влияющих на успех краудфандинговых проектов, таких как популярность темы, длительность кампании, общий объем сбора средств, частота и качество выкладываемого контента, обратная связь между разработчиками и заинтересованными лицами [3].

Краудфандинговые платформы могут использовать алгоритмы машинного обучения и анализа данных для определения успешных и неуспешных проектов, чтобы улучшить свои рекомендации и прогнозы. Анализ больших данных может предоставлять информацию, основанную как на прошлых кампаниях, тенденциях рынка, так и на исследовании потребностей клиентов и графиков успехов и неудач. Некоторые из используемых методов включают в себя:

- Анализ похожих кампаний. Данные собираются, обрабатываются и анализируются. Например, на основе данных о прошлых краудфандинговых кампаниях можно определить уровень риска новой кампании.

- Прогнозирование по датам. Модели машинного обучения могут быть созданы для прогнозирования будущих событий, таких как успех кампании в зависимости от даты запуска.

- Анализ социальных медиа. Использование алгоритмов анализа социальных медиа может помочь в определении популярности и распространенности кампании. Big Data может помочь оценить эффективность использования различных методов маркетинга и рекламы для привлечения инвесторов. Такие данные могут быть использованы для определения лучших стратегий для сбора средств, таких как оптимизация контента проекта [4].

- Сегментация. Использование больших данных вместе с сегментацией может помочь краудфандинговым платформам лучше понимать свою аудиторию, оптимизировать процесс финансирования проектов и улучшать пользовательский опыт. Мониторинг текущих тенденций. Анализ текущих тенденций и изменений на рынке может помочь в прогнозировании успеха кампании.

- Определение пути развития проекта: анализ Big Data помогает выявить, как проект может развиваться после завершения кампании краудфандинга. Данные об истории проектов, сопоставление продуктов и взаимодействие со сторонними платформами могут дать полезную информацию о том, какие возможности есть для расширения бизнеса.

Краудфандинг дал жизнь множеству проектов, а многие люди смогли реализовать свои мечты. Одним из ярких примеров является игровая консоль Ouya, которая стала огромным хитом, собрав 8,5 миллионов долларов всего за 29 дней на Kickstarter, использующем большие данные для точных прогнозов [5].

#### **Заключение.**

Таким образом Big Data также помогает кампаниям определить целевую аудиторию, создавать персонализированные стратегии маркетинга и оптимизировать процесс сбора средств.



Применение анализа позволяет улучшить результаты краудфандинговых кампаний, снизить риски и повысить шансы на достижение поставленных целей.

#### **Список литературы**

[1] Block J, Hornuf L, Moritz A. Which updates during an equity crowdfunding campaign increase crowd participation? *Small Business Economics*. 2018;50(1):3-27.

[2] Попов С.О. Источники и инструменты финансирования инноваций // *International scientific review of problems and prospects of modern science and education / collection of scientific articles. XLIV international correspondence scientific and practical conference (Boston. USA. April 24-25, 2018)*. Boston, 2018. С. 37-40.

[3] Аналитики Big Data. Типы аналитики = Big Data analytics and concepts. Types of analytics / И. В. Шилов [и др.] : VIII Международная научно-практическая конференция : сборник материалов VIII Международной научно-практической конференции, Минск, 11–12 мая 2022 года / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. А. Богущ [и др.]. – Минск, 2022. – С. 273–276.

[4] Корсунов, В. Ю. Роль решений, использующих Big Data, в управлении организациями = The role of big data solutions in the management of organizations / В. Ю. Корсунов, С. Н. Нестеренков : VIII Международная научно-практическая конференция : сборник материалов VIII Международной научно-практической конференции, Минск, 11–12 мая 2022 года / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. А. Богущ [и др.]. – Минск, 2022. – С. 257–260.

[5] Kaushik Paul: How is Big Data influencing crowdfunding, 2023

### **BIG DATA IN CROWDFUNDING AS A MEANS OF PREDICTING THE SUCCESS OF A CAMPAIGN**

***S.N. Barsukevich***

*4th grade student of the Faculty of  
Computer Systems and Networks,  
specialty CSaPT BSUIR*

***S.N. Nesterenkov***

*Dean of the Faculty of Computer  
Systems and Networks of BSUIR,  
PhD of Technical Sciences, Associate  
Professor*

***D.V. Nizovtsov***

*Head of Information Technology  
Department, ITD Assistant,  
Department of SFiT*

*Department of Computer Science*

*Faculty of Computer Systems and Networks*

*Belarusian State University of computer science and Radio Electronics, Republic of  
Belarus*

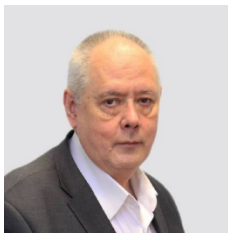
*E-mail: mx2002bsn0307@gmail.com*

**Abstract.** This article provides a brief overview of the field of crowdfunding, important aspects of the success of campaigns are analyzed, and also ways to use big data to predict a positive result.

**Keywords:** crowdfunding, data analysis, campaign success, monitoring are identified.

УДК 004.6

## ОБУЧЕНИЕ СПЕЦИАЛИСТОВ-МАРКЕТОЛОГОВ РАБОТЕ С ДАННЫМИ В СОВРЕМЕННЫХ УСЛОВИЯХ



**Н.В. Павлов**

Доцент СПбПУ, кандидат технических наук  
pavlov@kafedrapik.ru

### **Н.В. Павлов**

Окончил Ленинградский электротехнический институт (1980). Кандидат технических наук (1991). Доцент СПбПУ (с 1994), Doctor of Science (2016, Abo Academi University, Finland). Исследования и преподавание маркетинговых исследований, информационных систем, расширенного интеллекта.

**Аннотация.** Статья посвящена опыту обучения маркетологов основам работы с данными для принятия решений. Показан спектр компетенций, необходимых для работы в области маркетинга, основанного на данных. Часть из них предлагается осваивать с помощью сквозной задачи, на которой демонстрируются способы сбора данных, их очистки, хранения, аналитики и мониторинга ситуации во внешней среде. Обсуждаются особенности задачи. Предлагаются варианты дальнейшего ее расширения.

**Ключевые слова:** аналитика, маркетинговые исследования, обучение, цифровизация, данные.

В настоящее время цифровой маркетинг признается трендом [4], но, по мнению автора, он должен вскоре стать органической частью маркетинга в целом.

Будущее – за маркетингом, основанным на данных.

В литературе описаны не только перспективы [15], но и сложности и провалы внедрения систем хранения и аналитики данных [12, 18].

Это значит, что для работы в этой области требуется серьезная подготовка специалистов.

По мнению автора, процесс перехода к маркетингу, основанному на данных [14], в реальных быстро изменяющихся условиях должен происходить постепенно, в развитии. Этот же путь целесообразен для стартапов.

Данная статья посвящена опыту обучения маркетологов основам работы с данными для принятия решений.

На каждом этапе работы с данными имеются свои сложности.

Реальных данных в свободном доступе для целей обучения не так уж много [7], но и их ручной поиск и сбор весьма трудоемки.

Организации практически не делятся своими внутренними данными.

Даже электронные магазины и маркетплейсы принимают меры против сбора их данных ботами. Предоставляемые маркетплейсами для этого API требуют для использования специальных навыков.

Собираемые данные обычно не готовы для хранения, а тем более – для анализа [9, 13]. Их требуется очистить и сохранить.

Имеются различные подходы к вопросу хранения очищенных или неочищенных данных [16]. Очевидно, это определяется особенностями решаемых задач.

Для аналитики нужны извлеченные из хранилища и подготовленные данные. Подготовка данных для принятия решений – тоже специфическая область, в основном касающаяся визуализации.

Следует отметить, что для решения описанного круга задач интенсивно разрабатывается новый инструментарий, основанный на искусственном интеллекте.

Таким образом, рассматриваемая область весьма обширна. Поэтому требуется комплексное обучение процессу перехода к цифровому маркетингу.

Важная часть такого перехода – аналитика: подготовка информации для принятия решений.

Другая цель – дать системные знания, которые позволят свободно осваивать новый инструментарий.

В источниках последних лет довольно трудно разобраться, так как обычно авторы считают, что у читателей уже имеется большой багаж знаний, например, о нормальных формах отношений [5], CSS [6], Regexp [9].

Часто приходится отвлекаться на поиск описания базовых понятий.

Кроме того, в настоящее время большинство работ посвящено лишь продвижению в Интернет, воронке продаж [1, 3, 8, 11].

Для освоения навыков работы с реальными данными предлагается следующая задача, проверенная в рамках курса Информационные системы в маркетинге в течение трех лет.

Вариант 1. Для себя: купить лучший товар определенного вида (робот-пылесос, фотоаппарат, мобильный телефон и т.п.) в хорошем магазине по хорошей цене.

Вариант 2. Для организации: в какой магазин (с лучшим рейтингом, лучшим ассортиментом, лучшими условиями сотрудничества) лучше отдавать свой товар для продажи.

Оба варианта основаны на сборе данных

- о товарах, в первую очередь – их рейтингах;

- о магазинах;

- о прайс-листах этих магазинов.

Выполнение задания даст понятие о маркетинге, основанном на данных, об основных методах работы с данными.

Решение этой задачи по подготовленным методическим указаниям занимает 8...10 академических часов.

Сущность задачи состоит в

- сборе данных,

- их очистке,

- сохранении в виде системы связанных таблиц,

- выводе данных для принятия решения,

- решении дополнительных задач, например, многокритериального выбора.

Инструментарий для этих действий весьма разнообразен.

В настоящее время сбор данных производится с помощью бесплатных расширений браузера Google Chrome.

Схема данных представлена на рисунке.

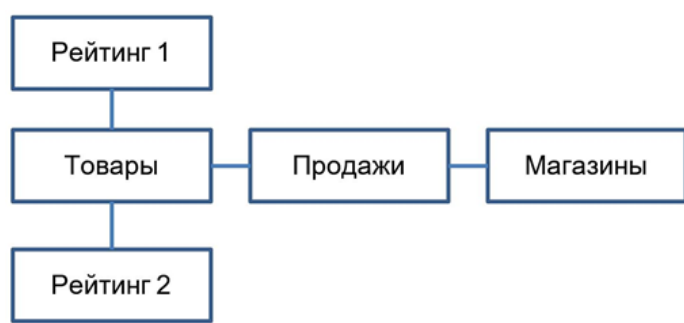


Рисунок 1. Схема данных

Примерный план занятий.

1. (2 академических часа.) Описание задачи. Сбор данных не менее, чем двух рейтингов заданного вида товара с помощью скрепера. Используется автоматический скрепер (Instant Data Scraper или автоматический режим NoCoding Data Scraper. Часто появляются новые, более удобные расширения). Результат – неочищенные данные, в которых, однако, просматриваются полезная информация: название продукта и место в рейтингах. Формат данных: .csv или .xlsx. Файлы показывают пример неочищенных данных.

2. (от 2 до 4 академических часов.) Создание т.н. рецептов сбора данных в расширении NoCoding Data Scraper. Его интеллектуальная поддержка упрощает эту работу. Задача состоит в том, чтобы в автоматическом режиме

- просмотреть список товаров на Яндекс.Маркете (сайт очень динамично меняется, сбор данных ботами постоянно усложняется); на Отзовике; на Shops-Prices.ru или аналогичном сайте;
- переходить по страницам этого списка;
- для каждого товара выбрать его название;
- для каждого товара перейти к списку магазинов, где он продается;
- для каждого магазина выбрать: цену продажи товара, название магазина, его рейтинг.

Для выполнения этих шагов требуется освоить механизм селекторов и управления работой скрепера. В частности, необходимо ввести задержки времени между действиями, чтобы имитировать работу человека-оператора. Процесс построения рецепта происходит достаточно наглядно: формируется графическая схема, имеется предварительный просмотр собираемых данных. Готовый рецепт обеспечивает переходы по страницам сайта.

Результат – большая таблица формата .csv или .xlsx. Но пока это – еще не полностью очищенные данные.

3. (2 часа.) Очистка данных с помощью языка M. Он включен в приложение MS Excel 2016 и более поздние версии, а также используется в бесплатном приложении Power BI Desktop. Результат – таблицы данных в соответствии с рисунком. Из неочищенных файлов удаляются лишние данные, преобразуется их формат, данные разделяются по содержанию. Результат – 5 таблиц, готовых к анализу.

4. (2 часа.) Связь таблиц в Excel или Power BI. Строится схема «снежинка» [14]. Тут возможен выбор. Либо продолжается прошлая работа в Power BI, тогда получается, что результат формируется из сохраненных неочищенных данных. Либо объединяются таблицы, взятые из сохраненных очищенных файлов.

Далее создается запрос, содержащий в табличном виде информацию для принятия решений: модель товара, его места рейтингах, перечень магазинов, где он продается, цены в них, рейтинги этих магазинов.

На этом основная часть задания завершается.

Дополнительные действия состоят в обоснованном выборе лучшего товара и места его покупки с учетом рейтингов товара, рейтингов магазинов и цен в магазинах.

В конце задания полезно провести обсуждение проделанной работы и особенностей функционирования задачи в реальных условиях. Главная идея состоит в том, что проделанная работа – основа регулярного контроля текущей ситуации на рынке. Повторное выполнение запросов позволит следить за изменением ассортимента и цен.

Обсуждается удобный способ хранения исторических и актуальных данных для облегчения мониторинга внешней среды.

Обсуждаются и варианты применения собранных данных для аналитики. Это может быть отслеживание динамики цен, стабильности ассортимента, постоянства присутствия товаров в магазинах, выявление зависимости уровня цен от рейтинга магазина и т.д. Это продемонстрирует развиваемость цифровизации не только в ширину, но и в глубину.

В перспективе видится расширение задачи. Например, включение условий доставки, сравнение товаров по отобраным по важности характеристикам.

В результате на практике демонстрируется:

- решение аналитической задачи;
- варианты хранения данных;
- реалистичный путь цифровизации маркетинговой аналитики.

Следует, однако, отметить и отрицательную сторону описанного задания. Многократное автоматическое обращение к сайту с помощью ботов не совсем этично. Для учебных целей был бы полезен кэш: однократное скачивание страниц сайта и дальнейшая работа с их копиями. Такую возможность предоставляет система Octoparse, но опыт работы с ней показал, что она более сложна и даже нестабильна! Возможно, появятся новые, более удобные версии этой системы.

Другим решением данной этической проблемы могли бы стать API (Application Programming Interface), специально предоставляемые торговыми организациями для сбора данных с их сайтов. Однако это потребует освоения языка программирования. Впрочем, делаются попытки облегчить использование API непрограммистами [2]. Это – наиболее актуальное направление дальнейшего развития темы.

Опыт преподавания показал, что решение поставленной задачи, особенно в варианте 1, вызывает у студентов определенный интерес, хотя они иногда и говорят, что они не программисты. Полученные знания и навыки могут найти и другие полезные применения. Например, аналогичная задача – скачивание расписания с сайта университета и преобразование его в удобную форму.

Таким образом, предлагаемое задание позволит на практике ознакомить обучающихся с идеями аналитики реальных маркетинговых данных.

#### **Список литературы**

- [1]. Миллер Д. Воронки продаж по методу StoryBrand. – М. Альпина Паблишер. – 2022. – 105 с.
- [2]. Морин А. Из API Яндекс.Маркета -> в Excel и Power BI: интеграция. // Блог проекта Директ-ПРО. – URL <http://directprobi.ru/blogs/api-yandex-market-excel-power-bi-connector-partner-integraciya-zakazy-orders-products-fbs/>.
- [3]. Мультиановская Д.В. Маркетплейсы: как научиться продавать. Wildberries, Ozon и другие. – М.: АСТ. – 2022. – 180 с.
- [4]. Назаров А. Д., Товмасын Н. Д. Цифровой маркетинг как современный тренд. // Московский экономический журнал. – 2020. - №6. – с. 866 – 872.
- [5]. Нестеров С. А. Базы данных. – М.: Юрайт. – 2023. – 230 с.
- [6]. Сидельников Г. Наглядный CSS. – СПб.: Питер. – 2022. – 224 с.
- [7]. Смирнов Н. BIG DATA&AI 2022: если не ввести оборот больших данных, то «все пропало»? // Директор. – URL <https://cio.osp.ru/articles/280422-BIG-DATAAI-2022-esli-ne-vvesti-oborot-bolshih-dannyh-to-vse-propalo>.
- [8]. Солодар М. А. Настольная книга Интернет-маркетолога. – М.: Эксмо. – 2021. – 432 с.
- [9]. Сушков А., Шпрингер Л. Очистка данных: кто их загрязняет и что аналитику с этим делать // Блог Я Практикума. – 2022. – URL <https://practicum.yandex.ru/blog/chto-takoe-ochistka-dannyh/>.
- [10]. Форта Б. Изучаем регулярные выражения. М.: Диалектика. – 2019. – 192 с.
- [11]. Югова А. Продвижение ВКонтакте. – М.: Времена. – 2022. – 320 с.
- [12]. Craig S. Why a Majority of Data Warehouse Projects Fail—and What Businesses Can Do. // Database Trends and Applications. – 2021. – URL <https://www.dbta.com/Editorial/Think-About-It/Why-a-Majority-of-Data-Warehouse-Projects-Fail-and-What-Businesses-Can-Do-145910.aspx>.
- [13]. Joshua. The Ultimate Excel Data Cleaning Guide in 2023. // Quick Table. – 2023. – URL <https://www.quicktable.io/blog/the-ultimate-excel-data-cleaning-guide-in-2023-cldjhpp2a115191logkimzfavw>.
- [14]. Monberg C. 2023 Predictions For Data-Driven Marketers. // Forbes. – URL <https://www.forbes.com/sites/forbestechcouncil/2022/12/19/2023-predictions-for-data-driven-marketers/?sh=6f9142526a43>.
- [15]. Moses B. What's Next for Data Engineering in 2023? 7 Predictions. – 2022. // Search Medium. – URL <https://towardsdatascience.com/whats-next-for-data-engineering-in-2023-7-predictions-b57e3c1bf2d3>.

[16]. Reis J., Housley M. Fundamentals of Data Engineering. – O'Reily. – 2022. – 544 p.

[17]. Taniar D., Wenny R. Data warehousing and Analytics. – Springer. – 2021. – 642 p.

[18]. Vartika. Difficulties of Implementing Data Warehouses. // GeeksForGeeks. – 2023. – URL <https://www.geeksforgeeks.org/difficulties-of-implementing-data-warehouses/>.

## **TRAINING OF MARKETING SPECIALISTS TO WORK WITH DATA IN MODERN CONDITIONS**

*N.V. Pavlov*

*Associate Professor of SPbPU,  
Candidate of Technical Sciences*

*Peter the Great St. Petersburg Polytechnic University*

*E-mail: pavlov@kafedrapik.ru*

**Abstract.** The article is devoted to the experience of teaching marketers the basics of working with data for decision-making. The range of competencies required to work in the field of data-based marketing is shown. Some of them are proposed to be mastered with the help of an end-to-end task, which demonstrates how to collect data, clean it, store it, analyze and monitor the situation in the external environment. The specifics of the problem are discussed. Options for its further expansion are offered.

**Keywords:** analytics, marketing research, training, digitalization, data.

УДК 004.021:004.75

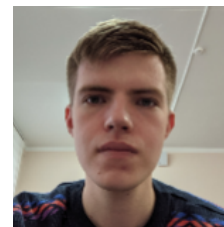
## ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ BIG DATA ПРИ ПОСТРОЕНИИ СИСТЕМЫ УПРАВЛЕНИЯ МОБИЛЬНЫМ РОБОТОМ



**М.А. Гергенсон**  
Студент факультета  
компьютерных систем и  
сетей БГУИР  
maratgerghenson@gmail.com



**С.Н. Нестеренков**  
Декан факультета  
компьютерных систем  
и сетей БГУИР,  
кандидат технических  
наук, доцент  
s.nesterenkov@bsuir.by



**И.С. Тарасюк**  
Ассистент кафедры  
электронных  
вычислительных машин  
БГУИР  
i.tarasiuk@bsuir.by

### **М.А. Гергенсон**

Студент четвертого года обучения по специальности "Вычислительные машины, системы и сети" БГУИР.

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцента кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

### **И.С. Тарасюк**

Окончил БГУИР в 2021 году по специальности "Вычислительные машины, системы и сети", магистрант первого года обучения по специальности "Системы и сети инфокоммуникаций" БГУИР.

**Аннотация.** В ходе разработки беспилотного мобильного робота возникла проблема представления и анализа большого потока неструктурированных данных, собранными подсистемами устройств. Для ее решения была создана система обработки данных, позволяющее развивать дальше разработанное устройство при помощи технологий анализа больших данных.

**Ключевые слова:** мобильный робот, подсистема, ROS, Hadoop, Spark.

### **Введение.**

Мобильный робот — это робот, способный передвигаться. Мобильная робототехника обычно считается подполем робототехники и информатики.

Мобильные роботы могут перемещаться в своей среде и не привязаны к одному физическому местоположению. Мобильные роботы могут быть «автономными» (AMR — автономный мобильный робот), что означает, что они способны осуществлять навигацию по неконтролируемой среде без необходимости использования физических или электромеханических устройств управления. В качестве альтернативы мобильные роботы могут полагаться на устройства управления, которые позволяют им перемещаться по заранее определенному маршруту навигации в относительно контролируемом пространстве (AGV — автономный управляемый автомобиль). Напротив, промышленные роботы обычно более или менее неподвижны, состоящие из сочлененного рычага (многосвязного манипулятора) и захватного узла (или концевой эффектора), прикрепленного к неподвижной поверхности.

Мобильные роботы стали более распространенными в коммерческих и промышленных условиях. Больницы используют автономные мобильные роботы для перемещения материалов в течение многих лет. В складах установлены мобильные роботизированные системы для эффективного перемещения материалов с полки для хранения в зоны выполнения заказов. Мобильные роботы также являются основным направлением текущих исследований, и почти у каждого крупного университета есть одна или несколько лабораторий, которые ориентированы на исследования мобильных роботов. Мобильные роботы также находятся в промышленных, военных и охранных системах. Внутренние роботы — это потребительские товары, в том числе развлекательные роботы и те, которые выполняют определенные домашние задачи, такие как пылесос или садоводство.

### **Структура мобильного робота.**

Беспилотные мобильные роботы способны передвигаться самостоятельно, благодаря специальному программному обеспечению (ПО) и сенсорам. ПО управляет работой всех систем робота: поворачиванием руля, сменой передач, газом и тормозом. Датчики (сенсоры) собирают информацию об окружающей обстановке, которая ложится в основу действий мобильного робота. Компоненты мобильного робота — это контроллер, программное обеспечение для управления, датчики и исполнительные механизмы. Контроллер обычно представляет собой микропроцессор, встроенный микроконтроллер или персональный компьютер (ПК). Программное обеспечение для мобильного управления может быть либо языком уровня сборки, либо языками высокого уровня, такими как C, C++, Pascal, Fortran или специальным программным обеспечением реального времени. Программное обеспечение беспилотного робота может включать компьютерное зрение и нейросети.

Большое число детекторов наделяет беспилотные роботы сверхчеловеческими способностями. Система, обеспечивающая беспилотное управление робота, включает в себя набор сенсоров, которые обеспечивают зрением машину, видеокамеры, радары, бортовое вычислительное устройство, которое обрабатывает сигнал, получаемый с этих датчиков. На этом устройстве формируется дорожная сцена, то есть устанавливаются объекты, которые находятся вокруг машины. Получается некоторая картинка, в центре которой находится робот, вокруг него находятся разного рода объекты: дома, другие автомобили, дорожные знаки, пешеходы, разметка, то есть всё, что нужно знать для того, чтобы мобильный робот мог в дальнейшем ориентироваться в пространстве: ехать прямо, перестроиться, притормозить, предпринять какие-либо действия, чтобы человека довести из точки А в точку Б, соблюдая все правила дорожного движения и, при этом, не попав в ДТП. Если один из детекторов обнаружил, что есть препятствия, дальше начинается оценка степени опасности этого препятствия, которое находится на траектории движения и оценивается время до столкновения. В полуавтоматическом режиме компьютер оповещает водителя, что впереди опасность и надо притормозить, если водитель не реагирует, то машина начинает тормозить самостоятельно. Обычно устанавливаемые датчики: лидары — дальномер оптического распознавания, радары, камеры, система глобального позиционирования (GPS, Глонасс), Датчики одометрии, гиростабилизатор.

### **Фреймворк Robot Operating System.**

Создание мобильного робота является командной работой, где параллельно ведется разработка нескольких систем, одной из которой и является автопилот, из чего произрастает вопрос о том, как обеспечить слаженную работу разрабатываемых систем. Для этого было решено использовать фреймворк ROS. Robot Operating System (ROS) — это гибкая платформа (фреймворк) для разработки программного обеспечения роботов. Это набор разнообразных инструментов, библиотек и определенных правил, целью которых является упрощение задач разработки ПО роботов. ROS была создана, чтобы



стимулировать совместную разработку программного обеспечения робототехники. Каждая отдельная команда может работать над одной конкретной задачей, но использование единой платформы, позволяет всему сообществу получить и использовать результат работы этой команды для своих проектов. Критериями, благодаря которым командой был выбран именно этот фреймворк, являются активность сообщества, наличие различных библиотек, расширяемость и простота использования. По этим критериям в данный момент равных ROS нет.

Все подсистемы в мобильном роботе связаны посредством ROS. Каждая из подсистем постоянно нуждается в информации, находящейся в другой, например, автопилот нуждается в данных с датчиков, модулей видеонаблюдения, геолокации и позиционирования, при этом всем все эти данные напрямую влияют на состояния автопилота и его поведение. Вся собранная информация о состоянии робота необходима для отслеживания системы, анализа работоспособности в не рассматриваемых на первоначальных этапах разработке ситуаций и сравнение результатов различных версий программы. Таким образом все накопленное придется хранить. ROS позволяет при помощи команды `rosv bag record` перехватывать все сообщения между узлами в указанном месте, однако в итоге мы получаем огромный пласт неструктурированных данных. Отладка такой системы и нахождение закономерностей – определенно тяжелая задача для выполнения вручную, да и данные продолжают неистово накапливаться во время использования.

#### Преобразование данных ROS в пригодные для анализа данные.

Предполагалось, что файлы пакета ROS должны быть преобразованы в более подходящий формат, прежде чем их можно будет обрабатывать параллельно с помощью таких инструментов, как Hadoop или Spark. Получается, что формат достаточно хорош для обработки с распределенной файловой системой типа HDFS. Для этого предназначен Hadoop `RosbagInputFormat` — разделяемый Hadoop `InputFormat` с открытым исходным кодом для формата файла `rosv bag`.

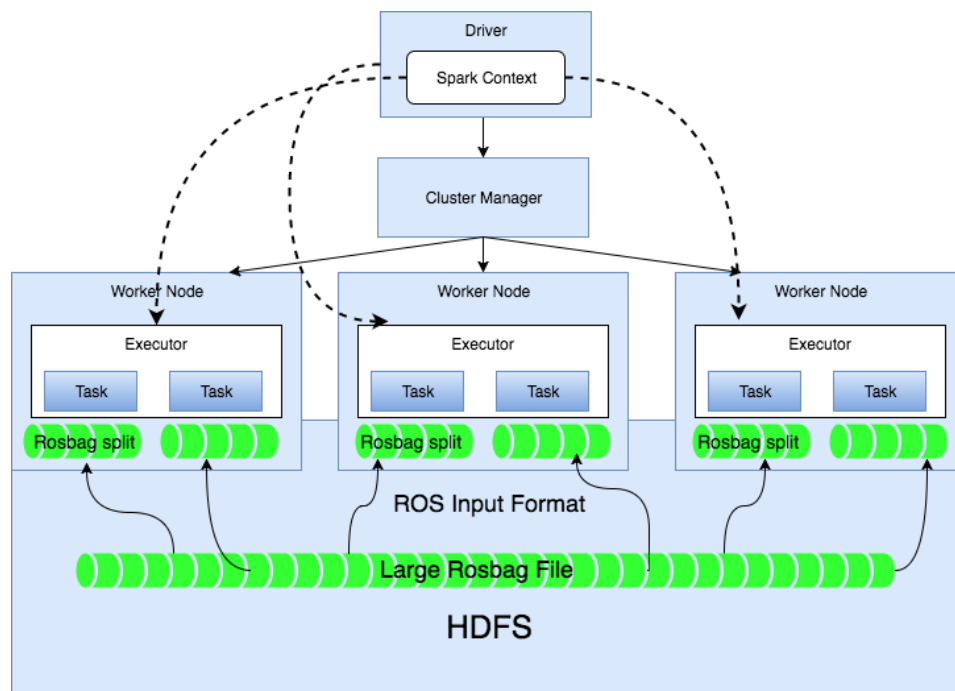


Рисунок 1. Пример разбиения файла Rosbag с Hadoop

Hadoop – это свободно распространяемый набор утилит, библиотек и фреймворк для разработки и выполнения распределённых программ, работающих на кластерах из сотен и тысяч узлов. Эта основополагающая технология хранения и обработки больших данных (Big Data) является проектом верхнего уровня фонда Apache Software Foundation. Изначально проект разработан на Java в рамках вычислительной парадигмы MapReduce, когда приложение разделяется на большое количество одинаковых элементарных заданий, которые выполняются на распределенных компьютерах (узлах) кластера и сводятся в единый результат.

Классический MapReduce, Apache компонент Hadoop для обработки данных, проводит вычисления в два этапа:

1 Map, когда главный узел кластера распределяет задачи по рабочим узлам.

2 Reduce, когда данные сворачиваются и передаются обратно на главный узел, формируя окончательный результат вычислений.

Пока все процессы этапа Map не закончатся, процессы Reduce не начнутся. При этом все операции проходят по циклу чтение-запись с жесткого диска. Это обуславливает задержки в обработке информации. Таким образом, технология MapReduce хорошо подходит для задач распределенных вычислений в пакетном режиме, но из-за задержек (latency) не может использоваться для потоковой обработки в режиме реального времени. Для решения этой проблемы был создан Apache Spark и другие Big Data фреймворки распределенной потоковой обработки (Storm, Samza, Flink).

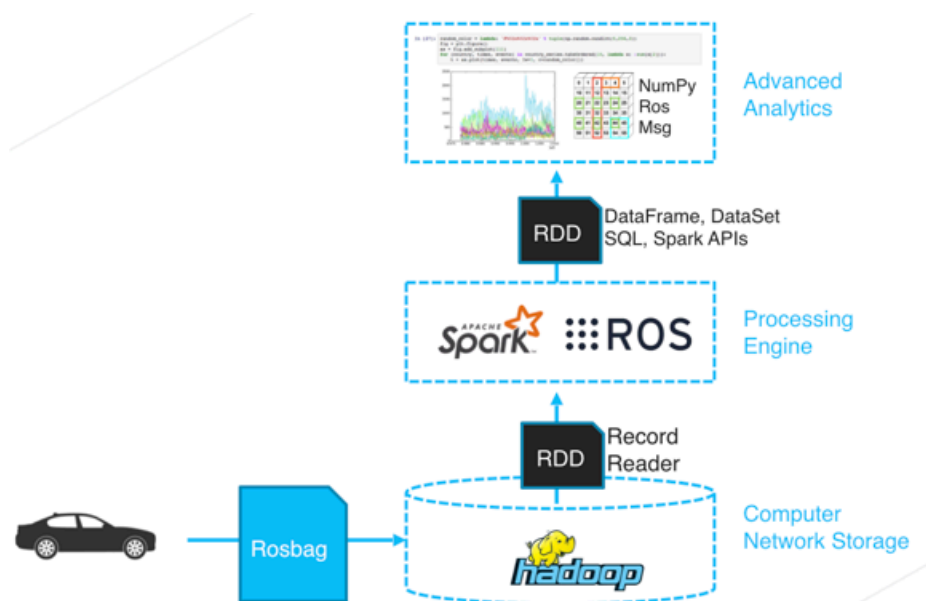


Рисунок 2. Итоговая структура обработки данных мобильного робота

В отличие от классического обработчика ядра Apache Hadoop с двухуровневой концепцией MapReduce на базе дискового хранилища, Spark использует специализированные примитивы для рекуррентной обработки в оперативной памяти. Благодаря этому многие вычислительные задачи реализуются в Spark значительно быстрее. Например, возможность многократного доступа к загруженным в память пользовательским данным позволяет эффективно работать с алгоритмами машинного обучения (Machine Learning).

Благодаря наличию разнопрофильных инструментов для аналитической обработки данных «на лету» (SQL, Streaming, MLLib, GraphX), Spark активно используется в системах

интернета вещей (Internet of Things, IoT) на стороне IoT-платформ, а также в различных бизнес-приложениях, в т.ч. на базе методов Machine Learning.

#### **Заключение.**

Возникшую при разработке беспилотного мобильного робота проблему представления и анализа данных работы подсистем машин, препятствующую дальнейшему развитию системы, удалось решить при помощи методов анализа больших данных на базе технологий Apache Hadoop и Spark.

В итоге была получена структура, позволяющая на ее базе реализовывать всевозможные виды систем обработки информации и управления, например, системы автопилотирования на базе методов машинного обучения, представления информации оператору и т.д.

#### **Список литературы**

- [1] Марц, Натан Большие данные. Принципы и практика построения масштабируемых систем обработки данных в реальном времени: моногр. / Натан Марц, Джеймс Уоррен. – М.: Вильямс, 2016. – 368 с.
- [2] Фуругян, Меран Алгоритмы планирования вычислений и синтеза систем реального времени / Меран Фуругян. – М.: LAP Lambert Academic Publishing, 2012. – 232 с.
- [3] Хетагуров, Я. А. Проектирование автоматизированных систем обработки информации и управления (АСОИУ). Учебник / Я.А. Хетагуров. – М.: Бином. Лаборатория знаний, 2015. – 240 с.
- [4] Lentin Joseph. Robot Operating System (ROS) for Absolute Beginners: Robotics Programming Made Easy / Lentin Joseph. — Apress, 2018.
- [5] Изучаем Spark. Молниеносный анализ данных / Х. Карау [и др.]. — ДМК Пресс, 2015. — 304 с.
- [6] Беляк, А. А. Анализ производительности технологии Hadoop / А. А. Беляк, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 19-20 мая 2021 года): / редкол.: В. А. Богуш [и др.]. – Минск: Бестпринт, 2021. – С. 343–346.
- [7] Одиноченко, М. И. Использование облачных сервисов для решения задач, связанных с применением технологии BIG DATA / М. И. Одиноченко, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VIII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 11-12 мая 2022 года): / редкол.: В. А. Богуш [и др.]. – Минск: Бестпринт, 2022. С. 252–256.

## **QUALITY EVALUATION OF INFORMATION TRANSFER IN A DISPATCHING SYSTEM BASED ON MQTT ARCHITECTURE**

***M.A. Herhenson***

*Student of the Faculty of Computer systems and networks of BSUIR*

***S.N. Nesterenkov***

*Dean of the Faculty of Computer systems and networks of BSUIR, PhD of Technical Sciences, Associate Professor*

***I.S. Tarasiuk***

*Assistant of the Department of Electronic Computers of BSUIR*

*Department of Electronic Computers*

*Faculty of Computer systems and networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: maratgergenson@gmail.com*

**Abstract.** During the development of an unmanned mobile robot, the problem of presenting and analyzing a large flow of unstructured data collected by device subsystems arose. To solve it, a data processing system was created that allows further development of the developed device using big data analysis technologies.

**Keywords:** mobile robot, subsystem, ROS, Hadoop, Spark.

УДК: 004.622

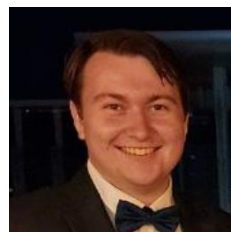
## РАЗВИТИЕ ИСПОЛЬЗОВАНИЯ НЕЙРОСЕТЕЙ ДЛЯ АНАЛИЗА ЭСТЕТИЧЕСКИХ ПРЕДПОЧТЕНИЙ С ПРИМЕНЕНИЕМ BIG DATA



**Д. И. Альховик**  
Студент 1 курса,  
кафедры  
информатики



**А.К. Горбачев**  
Студент 1 курса,  
кафедры информатики  
25350006@study.bsuir.by



**В.Д. Владымыцев**  
Ассистент кафедры  
информатики,  
инженер-программист  
ОИАСУ ЦИИР БГУИР  
v.vladymtsev@bsuir.by



**С.А. Мигалевич**  
Начальник  
ЦИИР, магистр  
технических  
наук

**Д.И. Альховик**

Студент 1 курса “Информатика и технологии программирования” БГУИР.

**А.К. Горбачев**

Студент 1 курса “Информатика и технологии программирования” БГУИР.

**В.Д. Владымыцев**

Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР

**С.А. Мигалевич**

Начальник ЦИИР, магистр технических наук

**Аннотация.** Данный доклад посвящен развитию использования нейросетей для анализа эстетических предпочтений с применением Big Data. В работе рассмотрены основные принципы работы нейронных сетей и способы их применения в анализе эстетических предпочтений. Также дан обзор использования Big Data в этой области, а также преимущества и недостатки такого подхода. В результате исследования было выяснено, что использование нейросетей в комбинации с Big Data позволяет добиться высокой точности в анализе эстетических предпочтений, что может быть полезно для многих отраслей, таких как дизайн, искусство, маркетинг и другие.

**Ключевые слова:** Нейросети, Эстетические предпочтения, Big Data, Анализ данных, Машинное обучение, Классификация, Регрессия.

### **Введение.**

В настоящее время Big Data, становится все более доступным и используются в различных сферах деятельности. В одной из таких сфер, связанных с искусством и дизайном, возможно применение нейронных сетей для анализа эстетических предпочтений. Это может быть полезно для таких отраслей, как мода, архитектура и музыка. С этой целью наша работа посвящена анализу развития использования нейросетей для анализа эстетических предпочтений на основе Big Data.

В данном докладе мы рассмотрим различные методы анализа эстетических предпочтений, обзор методов машинного обучения для работы с большими объемами данных и принципы функционирования нейронных сетей. Мы также проанализируем существующие исследования, использующие нейросети для анализа эстетических предпочтений. В результате мы проведем свое собственное исследование на основе данных, полученных из Big Data, и оценим эффективность использования нейронных сетей для анализа эстетических предпочтений.

В заключение нашего доклада мы обсудим результаты нашего исследования, его ограничения и перспективы развития данной области. Конечный результат нашего исследования позволит понять, насколько эффективно использование нейросетей для анализа эстетических предпочтений на основе больших объемов данных. Мы сможем выявить, насколько точно нейросети могут определить предпочтения людей в различных областях, таких как дизайн, музыка и мода. Наша работа поможет профессионалам в этих областях определять и прогнозировать новые тенденции и предпочтения своих клиентов, что может стать конкурентным преимуществом на рынке. Кроме того, мы также рассмотрим этические и социальные вопросы, связанные с использованием нейросетей для анализа предпочтений, и предложим возможные пути решения этих проблем.

#### **Основная часть.**

##### ***Методы анализа эстетических предпочтений***

Анализ эстетических предпочтений может быть выполнен с использованием различных методов, таких как опросы, экспертные оценки, сравнение изображений и многие другие. Однако, с развитием машинного обучения и доступностью больших объемов данных, появилась возможность использовать нейронные сети для анализа эстетических предпочтений.

##### ***Методы машинного обучения для анализа больших объемов данных***

Для анализа больших объемов данных можно использовать различные методы машинного обучения, такие как метод опорных векторов (SVM), случайный лес, нейронные сети и многие другие. Однако, при использовании нейронных сетей необходимо учитывать особенности их работы и правильно настраивать гиперпараметры.

##### ***Нейронные сети для анализа эстетических предпочтений***

Нейронные сети могут использоваться для анализа эстетических предпочтений, используя набор данных изображений, содержащий информацию об эстетических критериях, таких как симметрия, цветовая гамма, освещение и пропорции. Это позволяет создавать модели, которые могут оценивать качество изображения на основе этих критериев.

##### ***Применение Big Data для анализа эстетических предпочтений***

Big Data предоставляет большие объемы данных, которые можно использовать для обучения нейронных сетей для анализа эстетических предпочтений. Применение Big Data в анализе эстетических предпочтений позволяет использовать большое количество разнообразных изображений, что улучшает точность оценки эстетического качества и расширяет спектр приложений, включая обработку изображений для каталогов онлайн-магазинов, создание автоматических фотоальбомов, рекомендация фотографий и многие другие.

##### ***Примеры применения нейросетей для анализа эстетических предпочтений***

Самым популярным примером использования нейросетей является Midjourney. Midjourney использует собственную технологию глубокого обучения, позволяющую создавать изображения с высокой степенью детализации и реалистичности. Уже существует несколько готовых продуктов, таких как приложения для создания автоматических рекламных баннеров и обложек для книг на основе текстовых описаний.

Одним из примеров применения нейронных сетей для анализа эстетических предпочтений является работа, выполненная исследователями из университета Северной Каролины. Они разработали нейронную сеть, которая может оценивать эстетическое качество изображений на основе множества параметров, таких как цветовая гамма, контрастность, резкость и многие другие. Их исследование демонстрирует, что применение нейронных сетей позволяет получать более точные оценки качества изображений, чем традиционные методы анализа.



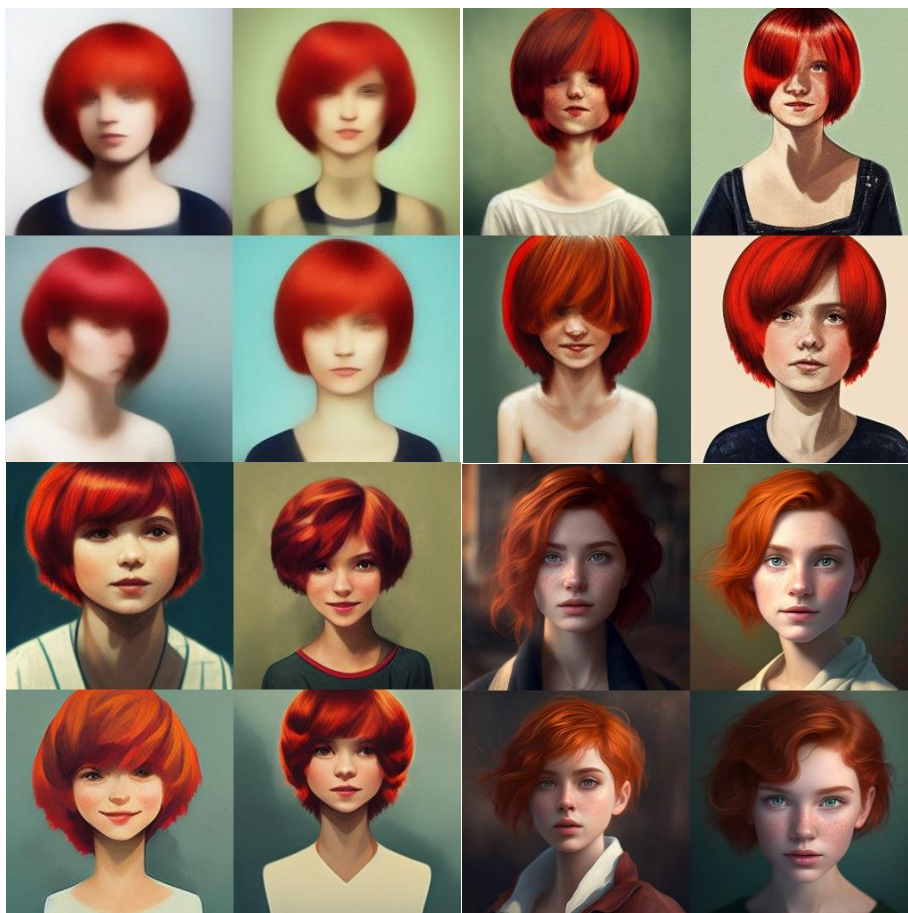


Рисунок 1. Результат улучшения работы нейросети Midjourney

Другой пример применения нейронных сетей для анализа эстетических предпочтений - это работа компании Google - Google Photos использует нейросеть для анализа фотографий и определения наиболее привлекательных изображений. Алгоритм анализирует такие характеристики, как резкость, цветовая гамма и композиция, чтобы выбрать лучшие фотографии из серии. Еще один пример применения нейронных сетей для анализа эстетических предпочтений - это работа китайских исследователей, которые создали нейронную сеть для оценки эстетического качества фотографий пейзажей. Они использовали сверточную нейронную сеть для извлечения признаков изображений и рекуррентную нейронную сеть для анализа эстетического качества. Результаты исследования показали, что их модель может давать высокую точность оценки эстетического качества фотографий пейзажей.

*Другие популярные нейросети связанные с работой в искусстве:*

DALL-E - это нейросеть, разработанная OpenAI, которая может генерировать изображения на основе текстового описания. Например, DALL-E может использоваться для создания изображений на основе эстетических характеристик, таких как цветовая гамма, форма и текстура объектов.

Art AI - это приложение для обработки изображений, которое использует нейросеть для создания художественных произведений в различных стилях и настроениях. Art AI анализирует характеристики изображения, такие как форма, цвет и текстура, а затем применяет алгоритмы, чтобы преобразовать его в стиле конкретного художника. Таким образом, Art AI может помочь пользователю лучше понять, какой стиль и настроение художественного произведения ему ближе по духу.

Artrendex - это сервис для анализа художественных произведений, который использует нейросеть для анализа характеристик искусства, таких как стиль, техника, эпоха и многие другие. Этот сервис может помочь пользователю лучше понять, какие художественные произведения более всего соответствуют его вкусам и предпочтениям, исходя из анализа тех характеристик, которые он наиболее ценит в искусстве.

EyeEm - это приложение для фотографов, которое использует нейросеть для анализа фотографий и определения их эстетического качества. Алгоритм анализирует композицию, цветовую гамму, фокусировку и другие характеристики фотографии, а затем присваивает ей рейтинг на основе их значимости. Этот рейтинг может помочь пользователю выбрать наиболее эстетически привлекательные фотографии для его потребностей.

Amazon использует нейросеть для анализа покупательских предпочтений и рекомендации товаров, основанных на эстетических характеристиках. Например, нейросеть может анализировать дизайн и цветовую гамму товаров, чтобы рекомендовать продукты, которые лучше соответствуют вкусам покупателя.

Adobe Sensei - это платформа искусственного интеллекта, которая использует нейросети для анализа эстетических характеристик изображений и оптимизации процесса обработки изображений. Например, Adobe Sensei может использоваться для автоматической коррекции цвета, конвертации фотографий в черно-белый формат или для создания оптимальной композиции фотографии.

Prisma - это приложение для обработки фотографий, которое использует нейросеть для преобразования фотографий в художественные произведения и имитации стилей различных художников. Нейросеть анализирует стиль и композицию фотографии, а затем применяет алгоритмы, чтобы преобразовать ее в стиле конкретного художника.

StyleMyPic - это приложение для обработки фотографий, которое использует нейросеть для оптимизации цвета, экспозиции и насыщенности фотографий. Алгоритм анализирует характеристики изображения и оптимизирует его для наилучшего эстетического вида.

Praxik - это стартап, который использует нейросеть для анализа эстетических характеристик веб-дизайна. Например, Praxik может анализировать цветовую гамму, шрифты, композицию и размещение элементов на веб-странице, чтобы определить, что будет наилучшим выбором для конкретного сайта.

Aiera - это платформа для анализа инвестиций, которая использует нейросеть для анализа эстетических характеристик отчетов компаний. Например, Aiera может анализировать цветовую гамму, шрифты, композицию и макет документов, чтобы определить, какие компании представляют больший интерес для инвесторов.

Artisto - это приложение для обработки видео, которое использует нейросеть для преобразования видео в художественные произведения в стиле различных художников. Нейросеть анализирует композицию, цвета и стиль видео, а затем применяет алгоритмы, чтобы преобразовать его в стиле конкретного художника.

Artomatix - это стартап, который использует нейросети для генерации 3D-моделей с помощью анализа эстетических характеристик. Например, Artomatix может анализировать текстуры и цвета объектов в реальном мире и использовать эти данные для создания новых 3D-моделей.

#### **Результаты исследования.**

Исследования в области использования нейросетей для анализа эстетических предпочтений с применением Big Data показали высокую эффективность такого подхода. В работе [3] было проведено исследование использования глубоких сверточных нейронных сетей для анализа эстетических предпочтений визуального контента, включая изображения. Авторы использовали набор данных из более чем 250 тысяч изображений, чтобы обучить свою модель. Исследователи отмечают, что использование больших данных (Big Data)

помогло значительно улучшить качество анализа. В частности, при обучении нейросетевой модели на большом количестве данных удалось достичь высокой точности предсказаний в задачах оценки эстетических качеств изображений. В работе [2] авторы предложили метод оценки качества цифровых изображений, который использует как визуальные, так и семантические характеристики изображения. Исследователи провели эксперименты на наборе данных из более чем 10 тысяч изображений и получили точность предсказаний более 80%. Исследования, упоминаемые в [1], показали, что нейросетевые модели могут использоваться для автоматической оценки качества фотографий, на которых присутствуют люди. Авторы использовали данные с социальных сетей и собрали набор из более чем 10 тысяч фотографий. Используя сверточную нейросеть, они добились точности предсказаний в 70%. Таким образом, результаты исследований подтверждают, что нейросетевые модели могут эффективно использоваться для анализа эстетических предпочтений визуального контента, в том числе изображений. Большие объемы данных (Big Data) помогают улучшить качество анализа и повысить точность предсказаний. Однако, несмотря на успехи в использовании нейросетей для анализа эстетических предпочтений, существуют и некоторые ограничения.

В работе [4] авторы обсуждают проблемы использования глубоких нейронных сетей в задаче оценки качества изображений. В частности, они указывают на то, что большой объем данных, необходимый для обучения нейросетевой модели, может стать проблемой при работе с ограниченным количеством данных. Кроме того, авторы отмечают, что для некоторых приложений, например, при оценке качества изображений в журналистике, может быть необходимо не только предсказание оценки качества, но и объяснение причин такой оценки.

Также в работе [5] авторы указывают на то, что использование нейросетей для анализа эстетических предпочтений может вызывать проблемы с этической точки зрения. Например, нейросеть может быть обучена распознавать определенные стереотипы, что может привести к неравенству и дискриминации. Кроме того, авторы отмечают, что использование нейросетей в задаче оценки качества изображений может быть неэффективно в случае, если изображение содержит информацию, которую невозможно выразить в виде числовых признаков. Тем не менее, несмотря на указанные ограничения, использование нейросетей для анализа эстетических предпочтений с применением Big Data остается актуальной и перспективной областью исследований.

#### **Обсуждение результатов.**

Как было показано в нашем исследовании, использование нейросетей в анализе эстетических предпочтений с применением больших данных является эффективным инструментом для создания более точных моделей прогнозирования. Тем не менее, остаются некоторые вызовы и проблемы, связанные с этой технологией, которые нужно рассмотреть. Во-первых, важно отметить, что интерпретация результатов моделирования глубоких нейронных сетей может быть сложной, так как эти модели обычно не обеспечивают никаких явных объяснений своих выводов [3]. Это может затруднить понимание причин, по которым определенные изображения оцениваются как эстетически привлекательные. Во-вторых, важно иметь в виду, что большие объемы данных, используемых в анализе эстетических предпочтений, могут привести к проблемам с доступностью данных и с конфиденциальностью [2]. Ограничения на использование и доступ к данным могут быть введены из-за различных причин, таких как конфиденциальность данных, ограничения прав доступа и нормативные требования. В этих случаях возможны другие методы работы с данными, такие как техника передачи обучения, которая позволяет использовать предварительно обученные модели для новых задач. Наконец, важно заметить, что создание точных моделей анализа эстетических предпочтений также зависит от корректного выбора функции потерь. Например,



использование L1-нормы для определения ошибок может привести к проблеме "размытости" изображений, что может негативно сказаться на точности прогнозирования [7].

Несмотря на эти вызовы и проблемы, использование нейросетей для анализа эстетических предпочтений является перспективным направлением исследований в области машинного обучения и компьютерного зрения. Однако для успешного развития этой технологии в будущем необходимо продолжать работу над улучшением точности моделей, а также улучшением их интерпретируемости и обеспечением конфиденциальности данных. Однако, несмотря на успехи и преимущества использования нейросетей для анализа эстетических предпочтений, есть и некоторые ограничения и проблемы, связанные с использованием Big Data. Во-первых, необходимо учитывать проблему «мусорных данных» (garbage in, garbage out). В случае недостаточной качественной подготовки данных, например, при использовании неадекватных метрик для оценки эстетических качеств изображений, может возникнуть проблема снижения точности модели. Во-вторых, нейросети могут создавать проблемы с точки зрения прозрачности и интерпретируемости. Это означает, что в случае использования нейросетей для анализа эстетических предпочтений, пользователи могут не иметь понимания того, как конкретная модель принимает решения и почему. В-третьих, использование Big Data может создавать проблемы с точки зрения конфиденциальности и защиты персональных данных. Сбор, хранение и обработка больших объемов данных могут привести к утечкам конфиденциальной информации.

Тем не менее, данные ограничения и проблемы не препятствуют дальнейшему развитию использования нейросетей для анализа эстетических предпочтений. С учетом того, что решения на основе нейросетей все больше становятся доступными и приобретают все большую популярность, в ближайшем будущем можно ожидать еще больших успехов в этой области. Еще одной проблемой является сложность интерпретации результатов, полученных с помощью нейросетей.

Также следует учитывать проблемы, связанные с качеством исходных данных. Нейросети могут быть чувствительны к шуму и ошибкам в данных, что может привести к низкой точности модели. Важно убедиться в том, что данные, используемые для обучения модели, являются достоверными и точными. Кроме того, следует учитывать этические вопросы, связанные с использованием нейросетей в анализе эстетических предпочтений. Например, использование нейросетей для прогнозирования предпочтений людей может вызывать опасения относительно конфиденциальности и защиты данных. Кроме того, некоторые люди могут возражать против использования алгоритмов машинного обучения для анализа их личных предпочтений.

#### **Заключение.**

В заключение можно отметить, что использование нейросетей для анализа эстетических предпочтений с применением Big Data представляет собой мощный инструмент, позволяющий обрабатывать большие объемы данных и учитывать сложные взаимосвязи между ними. Однако, для успешной реализации этой технологии необходимо учитывать ряд факторов, таких как объем и качество данных, выбор и обучение моделей, а также этические аспекты. Следует отметить, что нейросетевые модели не могут быть универсальным решением для всех задач, связанных с анализом эстетических предпочтений. Необходимо учитывать контекст и специфику конкретной области и принимать во внимание все ограничения и проблемы, связанные с использованием данной технологии. Кроме того, возможны проблемы с достоверностью результатов, особенно в случае недостаточного объема и/или нерепрезентативности данных, а также проблемы с интерпретируемостью и объяснимостью моделей. В целом, использование нейросетей для анализа эстетических предпочтений с применением Big Data является перспективным

направлением исследований. Однако, необходимо учитывать множество факторов, таких как объем и качество данных, выбор и обучение моделей, а также этические вопросы. Также следует учитывать специфику конкретной области и принимать во внимание все ограничения и проблемы, связанные с использованием данной технологии. В целом, использование нейросетей для анализа эстетических предпочтений с применением Big Data является мощным инструментом, но для его успешной реализации необходимо учитывать все ограничения и проблемы данного подхода и принимать меры для их решения.

### **Список литературы**

- [1] J. Lu, G. Wang, H. Deng, P. Moulin, “Aesthetic quality assessment of consumer photos with faces”, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1686-1693.
- [2] Y. Luo, X. Liu, “Aesthetic quality assessment of digital images with semantic and visual features”, Neurocomputing, vol. 311, 2018, pp. 231-239.
- [3] M. K. Khan, S. Hayat, S. A. Malik, S. A. Khan, “Deep convolutional neural networks for aesthetic analysis of images”, IEEE Access, vol. 5, 2017, pp. 8946-8958.
- [4] Y. Liu, J. Tang, Y. Liu, “Deep learning for image aesthetics: A survey”, Neurocomputing, vol. 266, 2017, pp. 321-332.
- [5] D. M. Hoi, R. W. K. Wong, “Towards robust and discriminative visual aesthetics modeling for photographic images”, in Proceedings of the ACM International Conference on Multimedia, 2011, pp. 877-880.
- [6] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? In Advances in neural information processing systems (pp. 3320-3328).
- [7] Molnar, C. (2019). Interpretable machine learning: A guide for making black box models explainable. Lulu.com.

## **DEVELOPMENT OF THE USE OF NEURAL NETWORKS FOR THE ANALYSIS OF AESTHETIC PREFERENCES USING BIG DATA**

**D. I. Alkhovik**  
*1st year student,  
Department of  
Computer Science*

**A.K. Gorbachev**  
*1st year student,  
Department of  
Computer Science*

**V.D. Vladymtsev**  
*Assistant of the  
Department of  
Informatics, software  
engineer of OIASU CIIR  
BSUIR*

**S.A. Migalevich**  
*Head of the Center  
for Informatization  
and Innovative  
Developments,  
Master of Technical  
Sciences  
migalevich@bsuir.by*

*Center for Informatization and Innovative Developments of the Belarusian State University of Informatics and Radioelectronics, Republic of Belarus  
Belarusian State University of Informatics and Radioelectronics, Republic of Belarus  
E-mail: 25350006@study.bsuir.by, v.vladymtsev@bsuir.by*

**Annotation.** This report is devoted to the development of the use of neural networks for the analysis of aesthetic preferences using Big Data. The paper considers the basic principles of neural networks and ways of their application in the analysis of aesthetic preferences. An overview of the use of Big Data in this area is also given, as well as the advantages and disadvantages of this approach. As a result of the study, it was found out that the use of neural networks in combination with Big Data makes it possible to achieve

**Keywords:** Data quality, Big data, Accuracy, Completeness, Timeliness, Reliability, Integrity, Security, Data validation, Data cleansing, Automated data update processes, Authentication, Access control, Encryption.

УДК 537.86:519.2

## ИСПОЛЬЗОВАНИЕ РАЗЛИЧНЫХ СЕМЕЙСТВ ВЕЙВЛЕТОВ В ЗАДАЧЕ СЛИЯНИЕ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ



**И.С. Куис**

Студент 4 курса специальности ЭВС,  
ФКСиС, БГУИР  
ivankuis11@gmail.com



**М.И. Вашкевич**

Доцент кафедры ЭВС  
БГУИР, доктор технических наук, доцент  
vashkevich@bsuir.by

### **И.С. Куис**

Студент четвертого курса БГУИР специальности «Электронные вычислительные средства». Область научных интересов: машинное обучение и цифровая обработка изображений.

### **М.И. Вашкевич**

Окончил БГУИР по специальности «Электронные вычислительные средства» в 2008 г. В 2013 г. защитил кандидатскую диссертацию по специальности 05.13.05 «Элементы и устройства вычислительной техники и систем управления». В 2022 защитил докторскую диссертацию по специальности 05.13.17 «Теоретические основы информатики». Работает доцентом кафедры ЭВС в БГУИР. Проводит научные исследования, связанные с цифровой обработкой аудиосигналов и изображений.

**Аннотация.** В работе представлены различные методы объединения медицинских изображений в области коэффициентов дискретного вейвлет-преобразования (ДВП). Также предложены варианты слияния, которые используют представление низкочастотной составляющей ДВП в области максимумов модуля вейвлет-преобразования, в его разреженном пространстве и с реконструкцией с использованием метода сопряженного градиента. Сравнение методов проводилось с использованием критерия нормированной взаимной информации для различных семейств вейвлетов.

**Ключевые слова:** Слияние изображений, ДВП, подавление немаксимумов, метод сопряженных градиентов, оценка качества слияния изображений.

### **Введение**

Важной задачей в медицинской практике является слияние (англ. *fusion*) изображений, полученных с использованием различного оборудования. Например, при планировании радиотерапии необходимо совмещать изображения различной модальности, т.е. полученных с датчиков различных физических полей. Так расчет дозы облучения основан на данных компьютерной томографии (КТ), а определение контуров опухоли часто лучше проводить на основе изображения, полученного методом магнитно-резонансной томографии (МРТ). Слияние изображений различной модальности позволяет лучше понять характер и особенности патологических изменений в организме [1].

Наиболее широко распространен подход к слиянию изображений на уровне пикселей (англ. *pixel-level image fusion*) [2]. Альтернативой является слияние на уровне признаков, в отличие от которого слияние на уровне пикселей напрямую комбинирует информацию из исходных изображений. В результате слияния должно получиться изображение, которое будет более информативным для визуального восприятия и компьютерной обработки.

В работе рассматривается применение многомасштабного преобразования для задачи слияния изображений. На рис. 1 показана общая схема применения многомасштабного преобразования для слияния изображений.

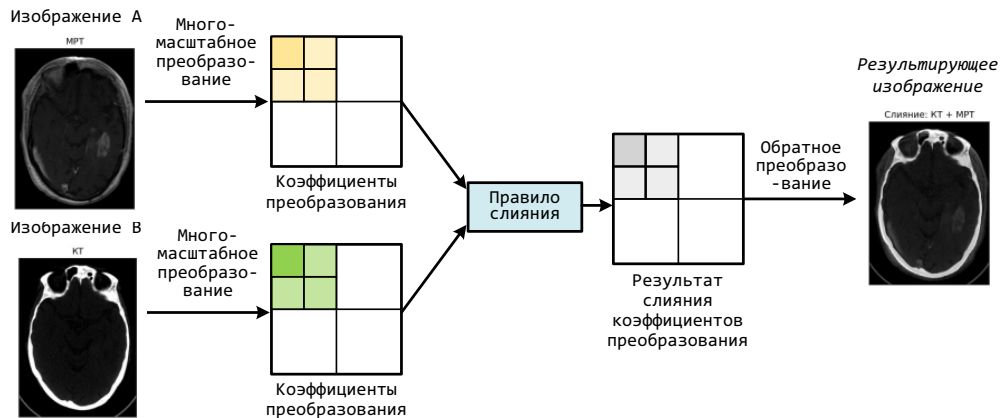


Рисунок 4. Схема слияния изображения

Как видно из рис. 1 слияния можно разделить на 3 этапа: 1) применение к исходным изображениям двумерного преобразования; 2) слияние изображений в пространстве коэффициентов; 3) восстановление результирующего изображения при помощи обратного преобразования. В работе рассматривается применение двумерного дискретного вейвлет-преобразования (ДВП), рассматривается вопрос повышения эффективности слияния изображений за счет представления НЧ компоненты изображений в разреженном пространстве максимумов модуля вейвлетов (англ. *wavelet modulus maxima*) [3], а также с восстановлением из пространства максимумов модуля вейвлетов с помощью метода сопряженных градиентов.

#### Применение двумерного ДВП в задаче слияния изображений.

Двумерное ДВП заключается в последовательном применении одномерного ДВП к строкам изображения, а затем к столбцам. В результате ДВП изображение разбивается на четыре квадранта – А, Н, V, D (рис. 2). Квадрант А – НЧ -коэффициенты, V – коэффициенты, полученные НЧ-фильтрацией по строкам и ВЧ-фильтрацией по столбцам; Н – коэффициенты, полученные ВЧ-фильтрацией по строкам и НЧ-фильтрацией по столбцам; D – ВЧ коэффициенты.

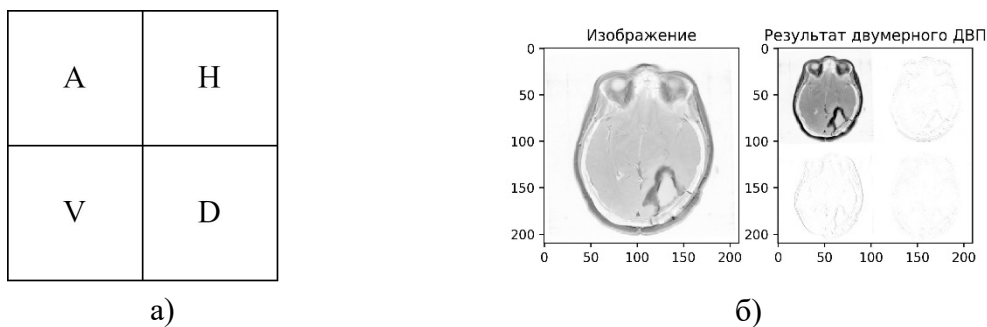


Рисунок 2. ДВП: а) Схема расположения коэффициентов; б) пример двумерного ДВП

Из примера двумерного ДВП, показанного на рис. 2,б можно видеть, что НЧ-коэффициенты А передают крупнозернистую структуру изображения, в то время, как ВЧ-коэффициенты V/H/D отвечают за границы и края. Общий принцип построения правил слияния заключается в обеспечении максимальной сохранности всех важных признаков, таких как границы и края. В [5] предложены следующие правила слияния:

$$1. \quad A_F = 1/2 \cdot (A_A + A_B), \quad (1)$$

$$2. \quad V_F = \max\{V_A, V_B\}, H_F = \max\{H_A, H_B\}, D_F = \max\{D_A, D_B\}, \quad (2)$$

где  $A_X, V_X, H_X, D_X$  – это соответствующие коэффициенты ДВП изображения  $X$ .

Согласно (1) НЧ-коэффициенты, отвечающие за крупные детали объединяются усреднением. Выражение (2) показывает, что слияние ВЧ-коэффициентов выполняется путем вычисления максимума. Это можно объяснить тем, что значения этих коэффициентов колеблются около нуля. Большие абсолютные значения ДВП в этих частотных полосах соответствуют резким изменениям интенсивности, т.е. важным деталям изображения (края, линии, границы). Поэтому хорошим правилом слияния является выбор наибольшего значения из двух коэффициентов ДВП в каждой точке. Обратным ДВП из коэффициентов  $A_F, V_F, H_F, D_F$  формируют результат слияния.

Далее будет рассмотрен метод представления изображения в пространстве максимумов модуля вейвлет-преобразования. Этот метод дает разреженное (англ. *sparse*) представление изображения. Мы используем его для обработки НЧ-коэффициентов ДВП  $A_A$  и  $A_B$ . В конечном итоге мы выполняем слияние НЧ коэффициентов по правилу (1), но только в пространстве максимумов модуля вейвлет-преобразования.

### Разреженное представление изображения в пространстве максимумов модуля вейвлет-преобразования (ММВП).

Переход в пространство ММВП выполняется путем сглаживания изображения с масштабированным ядром  $\theta(x, y)$ . Для вычислений используются вейвлеты  $\psi^{(x)}$  и  $\psi^{(y)}$ , которые являются частными производными сглаживающей функции  $\theta$ :

$$\psi^{(x)} = \frac{\partial \theta}{\partial x}, \psi^{(y)} = \frac{\partial \theta}{\partial y}.$$

Тогда можно ввести понятие двумерного вейвлет-преобразования  $W$  в точке  $(x, y)$  как:

$$\begin{aligned} Wf(x, y) &= \begin{pmatrix} W^{(x)}f(x, y) \\ W^{(y)}f(x, y) \end{pmatrix} = \begin{pmatrix} \psi^{(x)} * f(x, y) \\ \psi^{(y)} * f(x, y) \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x}(f * \theta)(x, y) \\ \frac{\partial}{\partial y}(f * \theta)(x, y) \end{pmatrix} \\ &= \nabla(f * \theta)(x, y). \end{aligned} \quad (3)$$

Выражение (3) соответствует градиенту изображения  $f(x, y)$ , модуль градиента в каждой точке вычисляется по формуле:

$$Mf(x, y) = \sqrt{|W^{(x)}f(x, y)|^2 + |W^{(y)}f(x, y)|^2},$$

а направление градиента относительно горизонтальной оси вычисляется как:

$$Af(x, y) = \arctg \left( \frac{W^{(y)}f(x, y)}{W^{(x)}f(x, y)} \right),$$

На рис. 3 показан пример применение выражений к реальному изображению. В качестве вейвлет-функции использовалось семейство Добеши (db1). Из приведенных результатов видно, что модуль градиента (см. рисунок 3,г) содержит границы областей изображения.

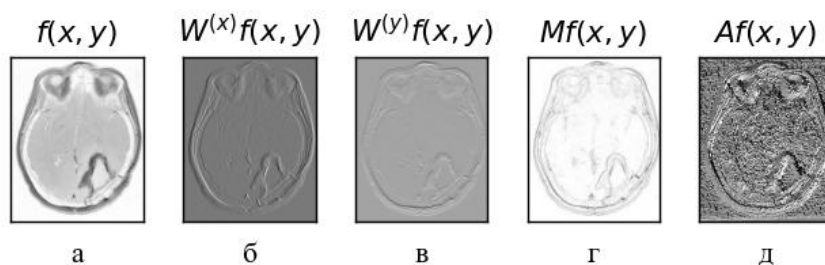


Рисунок 3. Пример вычисления ММВП: а) исходное изображение; б) градиент изображения по направлению  $x$ ; в) градиент изображения по направлению  $y$ ; г) модуль градиента; д) направление градиента

При слиянии двух изображений нам необходимо объединить все паттерны, которые присутствуют на обоих изображениях. Для этого нужно оставить только экстремумы поверхностей. Поиск максимумов состоит из двух этапов: 1) фильтрация изображения вейвлет-функцией по строкам и по столбцам (см. (3)) и 2) подавление не максимальных значений (англ. *NMS – non-maximal suppression*).

Метод NMS используется для определения границ на изображении. Пикселями границ являются те, в которых достигается локальный максимум градиента  $Mf(x, y)$  в направлении вектора градиента  $Af(x, y)$ . Алгоритм представляет собой итерацию по всем пикселям изображения, где для каждого пикселя выполняется следующие действия [3]: 1) квантование угла  $Af(x, y)$  до кратного  $45^\circ$  (для определения направления вектора градиента); 2) проверка соседних пикселей с текущим по направлению градиента. Если текущий пиксель больше соседних по модулю, то он остается. Иначе он приравнивается к нулю. Обозначим действие алгоритма подавления немаксимумов оператором:

$$Mf(x, y) = NMS\{Mf(x, y)\},$$

На рис. 4 показан пример подавления не максимальных значений.

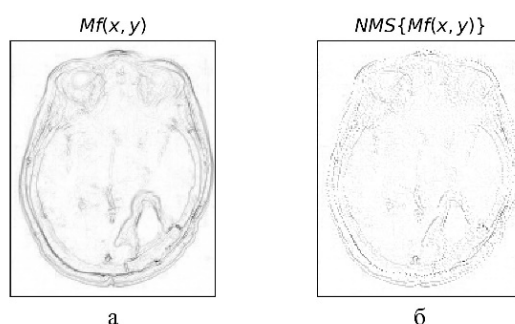


Рисунок 4. Пример подавления не максимальных значений

### Восстановление изображения из пространства ММВП.

Максимумы модуля вейвлет-преобразования (см. рис. 4,б) отражают свойства резких переходов и сингулярности на изображении. В [4] приводится метод, который восстанавливает аппроксимацию изображения по его вейвлет-максимумам. Ниже

приводится этот алгоритм реконструкции, который основан на теории фреймов (англ. *frame theory*).

В каждой точке ненулевой точке  $(u_p, v_p)$  пространства вейвлет-максимумов  $Mf(u_p, v_p)$  значение модуля и угла определяют две компоненты вейвлет-преобразования:

$$Mf(u_p, v_p) = \langle f, \psi_p^{(d)} \rangle, d \in \{x, y\},$$

где

$$\psi_p^{(d)} = \psi^{(d)}(x - u_p, y - v_p).$$

Алгоритм восстановления реконструирует изображение  $\hat{f}$  с минимальной нормой, такое, что

$$M\hat{f}(u_p, v_p) = \langle \hat{f}, \psi_p^{(d)} \rangle = \langle f, \psi_p^{(d)} \rangle.$$

Таким образом,  $\hat{f}$  – это ортогональная проекция  $f$  на замкнутое пространство  $V$ , генерируемое семейством вейвлетов  $\{\psi_p^{(x)}, \psi_p^{(y)}\}$ . Данное семейство является фреймом для  $V$ . В это случае  $f$  представляется, как

$$f = L^{-1}g, \tag{4}$$

где  $L$  – оператор выполняющий проекцию изображения в пространство  $V$ , а  $g$  – представление, изображения в пространстве  $V$ .

С другой стороны  $g$  можно определить используя  $\hat{f}$  следующим образом:

$$g = L\hat{f} = \sum_{d \in \{x, y\}} \sum_p \langle \hat{f}, \psi_p^{(d)} \rangle \psi_p^{(d)}. \tag{5}$$

На практике для получения  $\hat{f}$  вначале рассчитывают  $g$  по формуле (5), затем исходя из (4) вычисляют матрицу оператора  $L^{-1} = f g^{-1}$ . После чего исходя из (5) вычисляют  $\hat{f} = L^{-1}g$ . Следует отметить, что  $g$  часто является сингулярной матрицей, вследствие чего вместо  $g^{-1}$  приходится вычислять псевдообратную матрицу. На рис. 5 показан пример восстановления сигнала с использованием описанного алгоритма. Разница между исходным и восстановленным изображением находится на уровне  $1,4 \cdot 10^{-12}$ .

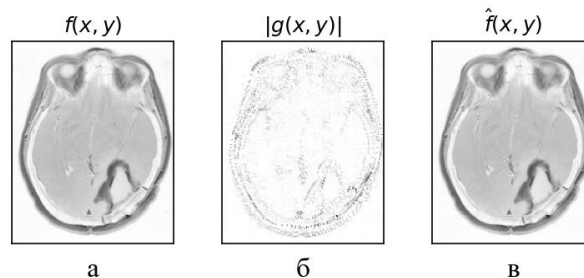


Рисунок 5. Восстановления изображения из разреженного пространства ММВП

### Восстановление изображений из пространства ММВП с использованием метода сопряженных градиентов.

Рассмотрим алгоритм восстановления НЧ-коэффициентов из пространства коэффициентов ДВП с помощью метода сопряженных градиентов. На вход алгоритм будет получать координаты изображения в пространстве ДВП  $g$ , полученные с помощью проекции оригинального изображения каркасным оператором  $A$  в пространство ДВП.

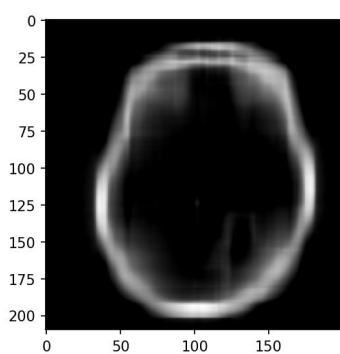
На каждой итерации алгоритма пересчитываются  $g$ , значение восстановленного изображения на  $f_k$  итерации, направление спуска  $d_k$ ,  $\beta_k$ , остаток  $r_k$ :

$$\begin{aligned} \alpha_k &= \frac{d_{k-1}^T r_{k-1}}{d_{k-1}^T A d_{k-1}} \\ f_k &= f_{k-1} + \alpha_k d_{k-1} \\ r_k &= g - A f_k \\ \beta_k &= \frac{-d_{k-1}^T r_k}{d_{k-1}^T A d_{k-1}} \\ d_k &= r_k + \beta_k d_{k-1} \end{aligned}$$

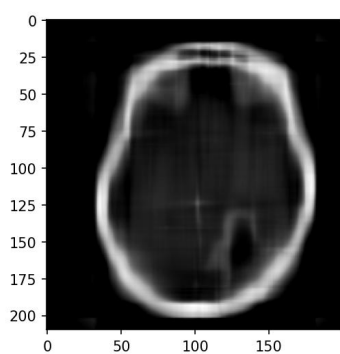
В качестве начальных значений используются следующие значения:

$$f_0 = 0, r_0 = g - A f_0, d_0 = r_0$$

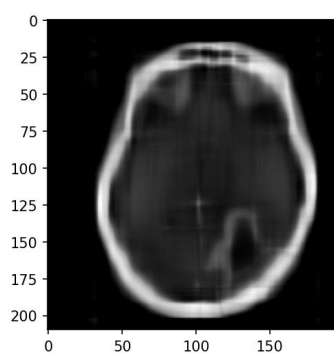
Главным преимуществом метода сопряженных градиентов по сравнению с обычным градиентным спуском является быстрая сходимость. Это происходит из-за того, что градиентный спуск на каждой итерации двигается только в направлении антиградиента функции, независимо от предыдущих направлений движения. Это может приводить к тому, что градиентный спуск "застревает" в ущельях или на плато функций, где он продолжает двигаться медленно в направлении минимума.



а)



б)



в)



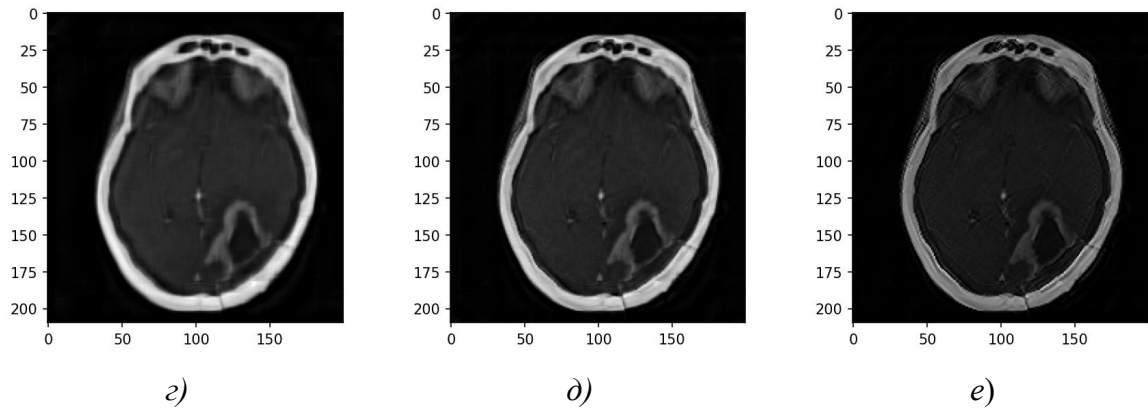


Рисунок 4. Результат восстановления: а) итерация 1; б) итерация 2; в) итерация 3; г) итерация 10; г) итерация 20; г) итерация 40;

В то время как в методе сопряженных градиентов направления спуска ортогональны друг другу относительно матрицы Гессе, что означает что они максимально некоррелированы и максимально информативными для оптимизации функции. Это и позволяет более эффективно двигаться в сторону минимума функции и быстрее достичь ее оптимального значения. Результаты восстановления изображения на разных итерациях алгоритма приведен на рис. 4.

#### Метод слияния изображений с использованием пространства ММВП.

Предлагаемый метод является улучшенной версией метода на основе двумерного ДВП с правилами слияния (1)–(2). Общая схема метода показана на рис. 5.

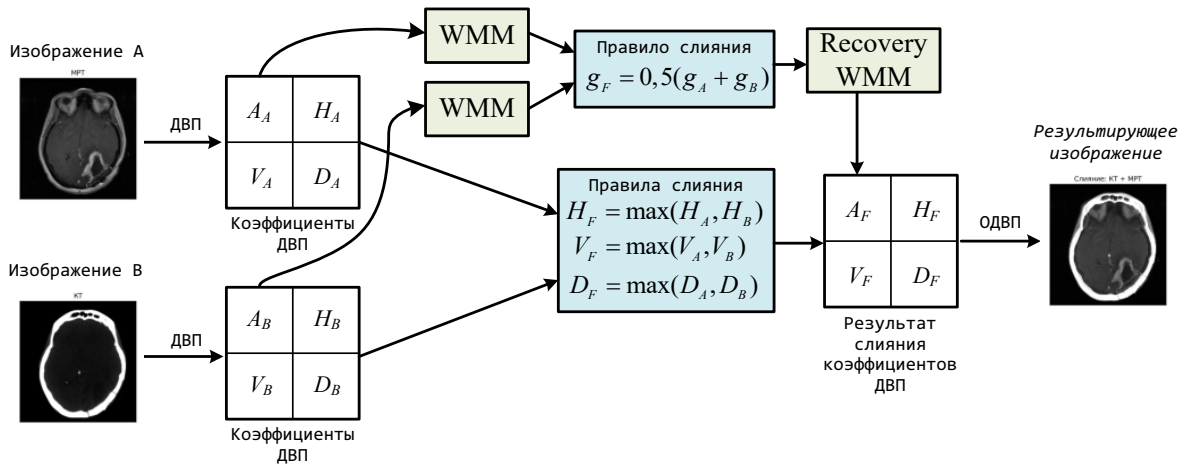


Рисунок 5. Слияние изображений с использованием пространства ММВП

Отличие предлагаемого метода от базового состоит в обработке НЧ-компонент ДВП объединяемых изображений. В (1) вместо НЧ-коэффициентов  $A_A$  и  $A_B$  объединяются их «образы» (5) в пространство вейвлет-максимумов. После чего выполняется восстановления НЧ-коэффициентов объединенного изображения  $A_F$  с использованием алгоритма, описанных в предыдущих разделах: с представлением НЧ коэффициентов изображений в разреженном пространстве максимумов модуля вейвлетов и восстановлением из пространства максимумов модуля вейвлетов с помощью метода сопряженных градиентов.

### Оценка качества слияния изображений.

Особенностью оценивания задач слияния изображений является эталонного изображения, которое должно получиться в результате слияния. Поэтому в качестве критерия качества слияния изображений при отсутствии опорного изображения в [6] предложено использовать понятие взаимной информации (ВИ). Данный критерий является мерой зависимости между двумя случайными величинами и измеряет количество информации, которое одна переменная содержит о другой переменной. Рассчитать взаимную информацию можно по следующей формуле:

$$MI(X, Y) = H(X) + H(Y) - H(X, Y),$$

где  $H(X)$  и  $H(Y)$  – распределение вероятности для изображений  $A$  и  $B$ , а  $H(X, Y)$  – совместная вероятность появления на изображении  $A$  пикселя с интенсивностью  $X$  и пикселя с интенсивностью  $Y$  на изображении  $B$ .

Главным недостатком данной метрики является чувствительность к масштабу. Например, если одна из переменных имеет большую энтропию, чем другая переменная, то большее значение их взаимных информаций этих переменных с третьей будет «перекошено» в сторону переменной с большей энтропией. Для решение этого недостатка введем понятие нормированной взаимной информации, которая будет нормироваться на максимум из энтропий двух переменных и вычисляться по следующей формуле:

$$NMI(X, Y) = \frac{MI(X, Y)}{\max(H(X), H(Y))}$$

Тогда можно ввести критерий  $M_F^{AB}$  для качества слияния двух изображений  $A$  и  $B$ , будет вычисляться как сумма взаимной информации изображений  $A$  и  $B$  с результатом слияния  $F$ :

$$M_F^{AB} = NMI(A, F) + NMI(B, F)$$

Таким образом  $M_F^{AB}$  показывает общее количество информации которое результат слияния  $F$  содержит об изображении  $A$  и  $B$ .

### База изображений и методика эксперимента.

Для оценки качества методов слияния изображений использовался набор из 30 пар изображений со снимками КТ и МРТ из онлайн ресурса The Whole Brain Atlas [7]. В ходе эксперимента 1) поочередно выбиралась пара изображений; 2) для каждой пары выполнялось слияние тремя методами; 3) вычислялась оценка слияния для результатов слияния.

### Результаты экспериментов с вейвлетами.

Выбор правильного семейства вейвлетов является важной частью алгоритма слияния изображений и может в значительной степени повлиять на конечный результат. Поэтому для методов, описанных выше были произведены эксперименты с разными вейвлетами: для вейвлетов каждого семейства оценивалось среднее значение критерия качества на всем наборе данных. Полученные результаты приведены в таблице 1.

В результате для алгоритмов с слиянием в коэффициентах ДВП и улучшение данного алгоритма с слиянием НЧ-коэффициентов в разреженном пространстве ММВП лучшее качество дало семейство койфлетов. Для алгоритма с слиянием НЧ компоненты в пространстве ММВП с восстановлением с помощью метода сопряженных градиентов

лучшее качество показали обратные биортогональные вейвлеты. Для разделимого двумерного ДВП на основе 8-канального банка фильтров среднее значение критерия было равно 0.8939.

Таблица 1. Результаты экспериментов с различными семействами вейвлетов

Семейство вейвлетов	ДВП	ДВП+ММВП + NMS	ДВП+ММВП + CG
Добеши	0.8975	1.0484	1.3519
Хаара	0.8673	0.9291	1.3800
Койфлеты	1.1133	1.1295	1.3872
Добеши-Мейер	0.9993	1.0725	1.4147
Симлеты	0.9348	1.0336	1.4558
Биортогональные	0.8242	0.955	1.5995
Обратные биортогональные	0.8063	0.9257	<b>1.6461</b>

### Заключение.

В работе выполнено сравнение трех методов слияния медицинских изображений, основанных на применении ДВП. Первый метод основан на использовании классического двухканального банка фильтров (ДВП) и одноуровневой декомпозиции изображения. Второй метод является улучшенным вариантом первого, в нем используется представление НЧ-коэффициентов ДВП в разреженном пространстве модулей максимумов вейвлет-преобразования (в работе он обозначен, как «ДВП+ММВП + NMS»). Третий метод также является улучшенной версией первого метода с представлением НЧ-коэффициентов в пространстве ММВП и последующим восстановлением с помощью метода сопряженных градиентов (обозначается как «ДВП+ММВП+CG»). В результате проведенных экспериментов с использованием 30 пар изображений КТ/МРТ установлено, что в среднем наилучшие метрики показывает подход «ДВП+ММВП+CG» с применением обратных биортогональных вейвлетов.

### Список литературы

- [1]. Хлесткин А. Ю. Методы формирования мультимодального изображения // Научные труды SWorld. – 2014. – Т. 4. – №. 4. – С. 14-18.
- [2]. Li S. et al. Pixel-level image fusion: A survey of the state of the art // Information Fusion. – 2017. – vol. 33. – P. 100-112.
- [3]. De Silva D. et al. Wavelet based edge feature enhancement for convolutional neural networks // Eleventh International Conference on Machine Vision (ICMV 2018). – SPIE, 2019. – P. 751-760.
- [4]. Qu G., Zhang D., Yan P. Medical image fusion by wavelet transform modulus maxima // Optics Express. – 2001. – Т. 9. – №. 4. – С. 184-190.
- [5]. Дворкович, В.П. Новый подход к использованию вейвлет-фильтров при обработке изображений / В.П. Дворкович, А.В. Гильманшин // Цифровая обработка сигналов – 2008 – № 1 – С. 37–42.
- [6]. Qu G., Zhang D., Yan P. Information measure for performance of image fusion // Electronics letters. – 2002. – vol. 38. – №. 7. – P. 313–315.
- [7]. The whole brain atlas [Электронный ресурс]. – 2023. – Режим доступа: <http://www.med.harvard.edu/aanlib/home.html>

## **USING DIFFERENT FAMILIES OF WAVELETS FOR TASK OF FUSION MEDICAL IMAGES BASED ON DISCRETE WAVELET TRANSFORM**

***I.S. Kuis***

*Student of faculty Computer systems and networks of BSUIR*

***M.I. Vashkevich***

*Associate Professor, Department of Computer Engineering of BSUIR, Doctor of Engineering, Associate Professor*

*Department of Information and Computer Systems Design*

*Faculty of Computer Engineering*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: alexvikt.minsk@gmail.com*

*Computer Engineering Department*

*Belarusian State University of Informatics and Radioelectronics, Belarus*

*E-mail: vashkevich}@bsuir.by*

**Abstract.** The paper presents various methods of combining medical images in the space of discrete wavelet transform coefficients (DWT). Fusion variants are also proposed that use the representation of the low-frequency component of the DWT in the region of the wavelet modulus maxima, in its sparse space and with reconstruction using the conjugate gradient method. The methods were compared using the normalized mutual information criterion for different families of wavelets.

**Keywords:** Image fusion, DWT, Non-maximal suppression, conjugate gradient method, image fusion quality estimation.

УДК 37.014

## ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ BIG DATA ДЛЯ ОТСЛЕЖИВАНИЯ И КОНТРОЛЯ УСПЕВАЕМОСТИ УЧАЩИХСЯ



**А.О. Чаплинский**  
Студент 4-го курса  
информатики и технологий  
программирования, факультета  
КСиС, БГУИР  
andreitipo@gmail.com



**С.Н. Нестеренков**  
Декан факультета  
компьютерных систем и  
сетей БГУИР, кандидат  
технических наук, доцент  
s.nesterenkov@bsuir.by



**И.Г. Скиба**  
Ведущий инженер-  
программист ОИТ,  
ассистент каф. ЭВС и  
ЭВМ  
i.skiba@bsuir.by

### **А.О. Чаплинский**

Студент 4-го курса информатики и технологий программирования, факультета компьютерных систем и сетей БГУИР.

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации"

### **И.Г. Скиба**

Окончила БГУИР в 2020 году по специальности "Вычислительные машины, системы и сети", магистрант первого года обучения по специальности "Компьютерная инженерия" БГУИР.

**Аннотация.** Данная статья рассматривает возможности использования технологии Big Data в образовании для отслеживания и контроля успеваемости учащихся. Описывается, как анализ больших объемов данных, сгенерированных в процессе обучения, может помочь выявить закономерности и тенденции в успеваемости учащихся, а также определить причины успехов и неудач. Обсуждаются практические преимущества применения этой технологии, в том числе возможность своевременного реагирования на проблемы и помощи учащимся в достижении лучших результатов.

**Ключевые слова:** использование технологии Big Data, отслеживание успеваемости, контроль успеваемости, образование, анализ данных, преподавание, учащиеся.

### **Введение.**

В последние годы технология Big Data все шире применяется в различных областях, включая образование. Одно из перспективных применений этой технологии - отслеживание и контроль успеваемости учащихся [1].

Анализ больших объемов данных, генерируемых в процессе обучения, может дать ценные выводы о достижениях студентов и помочь педагогам принимать обоснованные решения относительно образовательных стратегий и мероприятий. Идентифицируя паттерны и тенденции в успеваемости студентов, учителя и администраторы могут адаптировать свои методики, чтобы соответствовать индивидуальным потребностям учащихся и улучшить общие результаты обучения.

Одной из главных причин создания такой системы является необходимость обеспечения более эффективного и прозрачного образовательного процесса. При использовании данной системы учителя могут быстро и легко отслеживать успеваемость своих студентов, выставлять оценки и делать замечания. Родители же имеют возможность

узнавать о текущих успехах своих детей и оперативно реагировать на проблемы, если таковые возникают.

Одной из основных проблем текущих систем по отслеживанию успеваемости обучающихся является ограниченность возможностей анализа и интерпретации данных [2]. Нередко такие системы предоставляют лишь базовую статистику, которая может быть недостаточной для получения полной картины успеваемости студентов.

Другая проблема заключается в том, что такие системы не всегда учитывают индивидуальные потребности и способности каждого студента. Это может привести к некорректной оценке успеваемости, особенно если в анализе не учитываются некоторые факторы, такие как социально-экономический статус, уровень знаний и др.

#### **Актуальность.**

Во-первых, количество данных, генерируемых в процессе обучения, постоянно растет, и эффективный анализ этих данных может помочь преподавателям и родителям выявлять проблемы и помогать учащимся в достижении лучших результатов [3].

Во-вторых, использование технологии Big Data может повысить качество образования и успеваемость учащихся за счет более точной и своевременной оценки их успехов и проблем.

Наконец, в условиях быстро меняющейся образовательной среды и увеличивающихся требований к качеству образования, использование технологии Big Data становится необходимым для преподавателей и родителей, которые стремятся обеспечить наилучшие условия обучения для учащихся.

Цель использования технологии Big Data для отслеживания и контроля успеваемости учащихся состоит в улучшении образовательного процесса путем предоставления быстрого и удобного доступа к более полной и точной информации о текущей успеваемости студентов. Оно поможет учителям и родителям оценивать учебный прогресс учащихся, выявлять их сильные и слабые стороны, а также принимать соответствующие меры для улучшения качества обучения.

#### **Использование.**

Использование технологии Big Data для отслеживания и контроля успеваемости учащихся может предоставить ценную информацию об обучающихся и позволить улучшить процесс обучения [4].

Для начала, следует установить, какие данные необходимы для оценки успеваемости учащихся. Это могут быть результаты тестов, оценки, а также данные об участии в уроках и других мероприятиях.

Далее, эти данные могут быть обработаны с помощью алгоритмов машинного обучения, чтобы выделить паттерны и тренды, которые могут помочь выявить слабые места и сильные стороны студентов, а также предсказать их успех в будущем [5].

Технология Big Data также может помочь в организации персонализированного обучения, позволяя учителям настраивать программы и задания для каждого студента на основе его индивидуальных потребностей и уровня знаний.

Кроме того, анализ данных с использованием технологии Big Data может помочь учителям и администрации школы сделать более осознанные решения в отношении программ обучения, учебных планов и стратегий, основанных на реальных данных о успеваемости учащихся.

Одним из способов использования Big Data для отслеживания успеваемости является создание индивидуальных профилей учащихся, в которых будут содержаться данные о их академических достижениях, прогрессе и поведении. Эти профили могут быть использованы для выявления паттернов в успеваемости, что поможет учителям и администрации лучше понимать потребности и проблемы каждого учащегося.

Big Data также может помочь в создании персонализированных образовательных программ и улучшении процесса обучения. Например, анализ данных может показать, какие методы обучения работают лучше всего для конкретных групп учащихся, что позволит адаптировать учебный материал под индивидуальные потребности каждого студента [6].

Учителя могут использовать данные об успеваемости каждого студента, чтобы определить, какие темы нуждаются в дополнительной работе, а какие уже понятны. Big Data позволяет учителям создавать более эффективные программы обучения и учебные планы.

Анализ данных может также помочь в выявлении причин плохой успеваемости и разработки программ для их устранения. Big Data позволяет учителям и администрации школы принимать осознанные решения на основе данных.

Благодаря данному приложению учителя и родители могут легко отслеживать академический прогресс каждого студента, а также получать уведомления о любых изменениях в их успеваемости.

Это помогает сохранять постоянную связь между учителями, родителями и учащимися, что в свою очередь создает благоприятную образовательную среду и повышает качество обучения. Благодаря использованию современных технологий, таких как интерактивные графики и диаграммы, система сможет обеспечить быстрый и удобный доступ к информации об успеваемости студентов. Это поможет лучше оценивать прогресс каждого учащегося, а также выявлять их сильные и слабые стороны.

#### **Заключение.**

Технология Big Data предоставляет огромный потенциал для улучшения образования и повышения успеваемости учащихся. С ее помощью можно получить ценные данные об успеваемости студентов, выявить их сильные и слабые стороны, а также предоставить персонализированное обучение, что может значительно повысить качество образования и повысить успеваемость учащихся.

#### **Список литературы:**

- [1] The future learning environment, pedagogical and technological perspectives [Электронный ресурс]. – Режим доступа: <https://www.springeropen.com/collections/futurelearningenviron> - Дата доступа 23.03.2023.
- [2] REVIEW OF MONITORING TOOLS FOR E-LEARNING PLATFORMS, Ali Aloyayr and Atta Badii [Электронный ресурс]. – Режим доступа: <https://arxiv.org/ftp/arxiv/papers/1407/1407.2437.pdf> - Дата доступа 23.03.2023.
- [3] Нестеренков, С.Н. Проблематика и актуальность информационной системы учета успеваемости студентов / С.Н. Нестеренков, Н.В. Ющенко, А.Д. Радкевич // Проблемы повышения эффективности образовательного процесса на базе информационных технологий. 2019. - С. 95-98.
- [4] Big Data in Education. A Bibliometric Review, José-Antonio Marín-Marín, Jesús López-Belmonte, Juan-Miguel Fernández-Campoy and José-María Romero-Rodríguez [Электронный ресурс]. – Режим доступа: <https://www.mdpi.com/2076-0760/8/8/223> - Дата доступа 27.03.2023.
- [5] Big Educational Data & Analytics: Survey, Architecture and Challenges [Электронный ресурс]. – Режим доступа: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9093868> - Дата доступа 27.03.2023.
- [6] Нестеренков, С.Н. Функциональная модель процедур планирования и управления образовательным процессом как основа построения информационной среды учреждения высшего образования / С.Н. Нестеренков, Н.В. Лапицкая // Вести Института современных знаний. - 2018. - N 1. - С. 97-105.

## **USING BIG DATA TECHNOLOGY TO TRACK AND MONITOR STUDENT PROGRESS**

***A.O. Chaplinsky***

*4th year student of Informatics and programming technologies,*

***S.N. Nesterenkov***

*Dean of the Faculty of Computer Systems and Networks, Candidate of Technical Sciences, Associate Professor*

***I.G. Skiba***

*Lead Software Engineer, OIT assistant, Dept. EC assistant, Dept. ECS*

*Department of Computer Systems and Networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: andreitipo@gmail.com*

**Abstract.** The use of Big Data technology in education for tracking and monitoring student performance is a promising tool for improving the quality of education and increasing student achievement. The analysis of large amounts of data generated during the learning process can help to identify patterns and trends in student achievement, as well as determine the causes of success and failure. This, in turn, allows teachers and parents to timely respond to problems and help students achieve better results. However, to fully leverage the potential of Big Data technology, ethical and legal aspects must be considered, and data security and privacy must be ensured. The use of Big Data technology in education has the potential to revolutionize the way we approach teaching and learning, but it also raises important ethical and legal concerns.

**Keywords:** use of Big Data technology, progress tracking, progress control, education, data analysis, teaching, students.

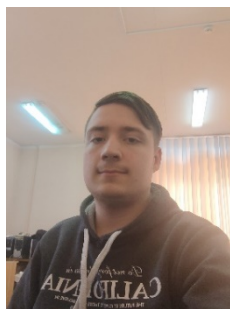


УДК 004.021:004.75

## АНАЛИЗ BIG DATA В ВЕБ-ПРИЛОЖЕНИЯХ ДЛЯ НАСТОЛЬНЫХ ИГР



**Г.А. Трофимук**  
Студент БГУИР  
gleb.trofimuk@gmail.com



**П.А. Архиреев**  
Студент БГУИР  
arhir.13.07.2002@gmail.com



**А.Н. Марков**  
Старший преподаватель,  
магистр технических наук,  
заместитель начальника  
Центра информатизации и  
инновационных разработок  
БГУИР  
a.n.markov@bsuir.by

**П.А. Архиреев**  
Студент кафедры информатики БГУИР.

**Г.А. Трофимук**  
Студент кафедры информатики БГУИР.

**А.Н. Марков**  
Магистр технических наук, старший преподаватель кафедры ПОИТ, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Показаны способы использования анализа Big Data в веб-приложения для настольных игр. Приведены примеры готовых решений и перспективных идей веб-приложений, основанных на анализе больших данных и связанных с настольными играми. Также для анализа больших данных веб-приложения, рассматривается распространённый фреймворк для работы с Big Data под названием Apache Storm.

**Ключевые слова:** Big Data, настольные игры, веб-приложения, Apache Storm.

### **Введение.**

Современные настольные игры не только сохраняют традиции прошлого, но и приспособляются к современным технологиям. Веб-приложения для настольных игр позволяют игрокам наслаждаться любимыми играми в онлайн формате. Одним из ключевых компонентов таких приложений является сбор и анализ больших объемов данных (Big Data), которые позволяют создавать индивидуальный игровой опыт для каждого игрока. В данной статье рассмотрены возможности использования Big Data в веб-приложениях для настольных игр и их влияние на улучшение игровой механики и игрового опыта.

### **Актуальность.**

С ростом количества данных, которые генерируются веб-приложениями, анализ больших данных становится все более важным для оптимизации пользовательского опыта и увеличения прибыли. Применение анализа больших данных в настольных играх также может помочь разработчикам понять, как улучшить игровой процесс, повысить вовлеченность игроков и увеличить доходы. В целом, материал представляет важные и

актуальные вопросы, которые относятся к анализу больших данных в веб-приложениях для настольных игр.

### Определение Big Data.

В настоящее время разрозненные и динамические объемы данных создаются инструментами, машинами и, что наиболее важно, людьми. Такой массивный пакет данных требует масштабируемых, инновационных, улучшенных и новых технологий для аналитики, хостинга и сбора. Технология больших данных способна обрабатывать собранные данные, чтобы предлагать предприятиям в режиме реального времени обширную информацию о пользователях, управлении производительностью, производительности, рисках и стоимости акционеров [1]. Скорость имеет решающее значение в этом мире больших данных. Традиционная аналитика обычно уделяет внимание анализу исторических данных, а с другой стороны, использование аналитики больших данных предполагает использование данных в реальном времени. Некоторые ведущие компании мира бизнеса, такие как Netflix, Capital One, Starbucks и Amazon, использовали большие данные для повышения эффективности своей работы.

### Создание идеального игрового бота.

В мире игр компьютерные боты стали неотъемлемой частью битвы за победу. Игроки используют их как помощников, а разработчики внедряют в игру, чтобы продемонстрировать свои технологические возможности. И в последние годы всё чаще игровые боты строятся на анализе большого количества матчей для выстраивания подходящей стратегии на каждый матч. Одним из успешного примера является шахматных движок Stockfish.

Stockfish – это один из самых сильных шахматных движков в мире. Он основан на открытом исходном коде, что позволяет разработчикам по всему миру вносить свой вклад в его развитие. В настоящее время Stockfish имеет рейтинг Эло около 3500, что делает его одним из наиболее мощных шахматных программ на рынке. Stockfish используется для анализа партий, критических ситуаций и многое другое. Компания Chess.com использует Stockfish как движок для своей онлайн-шахматной платформы.

В августе 2020 под влиянием бурного развития движков на основе ИИ Stockfish претерпел существенные изменения в архитектуре и стал использовать нейронные сети NNUE, оптимизированные под CPU и алгоритм поиска Stockfish наряду с традиционной оценочной функцией движка. Stockfish из классического движка стал гибридным, использующим оценочную функцию на базе нейронных сетей (есть возможность отключить использование нейронных сетей в настройках) в материально сбалансированных позициях и традиционную в остальных. Это изменение дало существенное усиление в силе игры [2].

На сегодняшний день Stockfish является самым успешным шахматным движком по рейтингу Эло:

Rank	Name	Rating			Score	Average Opponent	Draws	Games	LOS
		Elo	+	-					
1	Stockfish 15.1 64-bit 4CPU	3536	+15	-15	65.0%	-85.0	69.4%	1172	71.4%
2	Dragon by Komodo 3.2 64-bit 4CPU	3530	+16	-16	63.2%	-73.3	73.2%	1072	99.7%
3	Ethereal 14.00 64-bit 4CPU	3497	+19	-19	57.2%	-38.5	81.9%	690	69.4%
4	Berserk 11.1 64-bit 4CPU	3490	+19	-19	57.0%	-38.2	79.0%	690	88.1%
5	Koivisto 9.0 64-bit 4CPU	3475	+17	-17	55.1%	-27.8	82.0%	852	55.8%
6	Revenge 3.0 64-bit 4CPU	3474	+13	-13	54.1%	-23.4	80.6%	1604	55.3%
7	RubiChess 20221120 64-bit 4CPU	3472	+14	-14	54.5%	-24.3	80.6%	1373	59.6%
8	Igel 3.4.0 64-bit 4CPU	3470	+18	-18	54.0%	-22.1	83.4%	764	72.7%
9	SlowChess Blitz 2.9 64-bit 4CPU	3463	+12	-12	54.0%	-24.0	80.0%	1748	66.3%
10	Rebel 16.2 64-bit 4CPU	3458	+19	-19	52.4%	-12.5	83.3%	672	64.7%

Рисунок 1. Рейтинг Эло на момент 01.04.2023 [2]

Теперь разберем как работает Stockfish, который базируется на теории игр и использует обработку больших данных.

Шахматы — это игра, которая следует логическим концепциям, что делает ее детерминированной игрой. Это означает, что можно определить оценку позиции, не просматривая все возможные конечные состояния игры, что было бы практически невозможно из-за огромного количества узлов. Несмотря на свою детерминированность, в шахматах все же присутствует элемент случайности. В результате шахматные движки должны использовать вероятностный подход, чтобы увеличить свои шансы на победу. Stockfish присваивает каждой шахматной позиции десятичное значение, указывающее вероятность того, что каждый игрок выиграет или добьется ничьей. Этот подход отличается от теоретических оценок, которые дают только один из трех результатов: ничья, вынужденный мат через определенное количество ходов для белых или черных, или оптимальная стратегия по Парето (ни одна из сторон не имеет преимущества). Единицей оценки являются санти-пешки (centi-pawns (не знаю, как правильно перевести) [3]), величина которых больше 1,5 является значимой и указывает на высокие шансы на победу одной из сторон. Напротив, величина больше 4 указывает на полное доминирование. Точно так же смешанное равновесие Нэша включает вероятностные стратегии, а не чистые, и игроки выбирают ходы на основе связанных с ними вероятностей. Однако такие машины, как Stockfish, всегда выбирают ход с наибольшей вероятностью успеха.

Stockfish использует статическую функцию оценки или функцию, которая вычисляет оценочную оценку позиции с помощью эвристики (алгоритм, созданный вручную, предназначенный для соответствия любой произвольной шахматной позиции), который разрабатывался вручную учеными-компьютерщиками и гроссмейстерами более десяти лет [3]. Шахматный движок Stockfish анализирует Big Data, которая в нашем случае содержит данные о бесчисленном количестве проведенных шахматных партий (например, набор сыгранных партий с ресурса Chess.com [4]), чтобы предсказать оптимальный ход в любой ситуации. Это достигается за счет использования алгоритмов машинного обучения. Движок использует несколько методов Big Data, включая хранение больших объемов данных, параллельную обработку и распределение вычислительной нагрузки. Stockfish хранит информацию о миллионах ходов и позиций, чтобы иметь доступ к ним в любой момент и использовать эти знания для принятия решений. Stockfish также использует параллельную обработку, чтобы ускорить вычисления. Для этого он распределяет задачи на несколько процессоров и обрабатывает их одновременно. Это позволяет сократить время на обработку данных и ускорить принятие решений. Таким образом, вместо того, чтобы оценивать весь путь до мата, Stockfish просто смотрит так далеко, как может (обычно на 30+ шагов в глубину), и использует эвристику, основанную на стадии игры, чтобы присвоить значение тому, насколько «хорошо» действительно основан ход, вне полученного положения.

### **Генерация игрового персонажа для настольной ролевой игры.**

На сегодняшний день рынок настольных ролевых игр продолжает развиваться и набирать аудиторию. В настольных ролевых играх стоит выделить одного из самых успешных представителей данного жанра – Dungeons & Dragons (D&D, DnD). По состоянию на январь 2023 года в Dungeons & Dragons насчитывается около 50 миллионов игроков по всему миру, большинство из которых являются случайными, а не активными или одновременными. D&D Online, выпущенная в 2012 году, насчитывает в общей сложности 462 387 игроков или подписчиков и примерно 21 963 активных игрока каждый день [5].

В силу специфики игры, подход к ней у каждого игрока может быть совершенно разным и в основном ориентирован на личные предпочтения игрока, который позже принимает свою форму в виде персонажа и его действий. Если взять в расчет веб-приложение, которое позволяет проводить партии онлайн, то возможно фиксировать весь тот набор информации, который несет в себе игровая сессия. Полученный набор будет

представлять из себя неструктурированную информацию большого объема, анализ которой может быть использован для различных целей. Далее рассмотрим фреймворк для обработки больших данных Apache Storm.

### **Apache Storm.**

Apache Storm – это распределенная система обработки потоков данных, которая позволяет обрабатывать высокий объем потоков данных в реальном времени. Apache Storm имеет множество вариантов использования: аналитика в реальном времени, машинное обучение в режиме онлайн, непрерывные вычисления, распределенный RPC, ETL и многое другое. Apache Storm работает быстро: тест показал, что он обрабатывает более миллиона кортежей в секунду на узел. Он масштабируемый, отказоустойчивый, гарантирует обработку ваших данных, прост в настройке и эксплуатации [6].

Apache Storm имеет следующие преимущества: 1. Высокая производительность: Storm может обрабатывать миллионы записей в секунду на каждом узле кластера. 2. Масштабируемость: Storm легко масштабируется горизонтально, позволяя обрабатывать большие объемы данных. 3. Надежность: Storm обеспечивает гарантированную доставку данных и обнаружение сбоев, что делает его очень надежным. 4. Гибкость: Storm может быть использован для различных задач обработки потоков данных, включая обработку событий, машинное обучение, анализ данных и т.д.

Однако, у Apache Storm есть и некоторые недостатки: 1. Сложность: Storm имеет довольно высокий порог входа, и его настройка может быть сложной для новичков. 2. Необходимость управления: Storm требует управления кластером, что может быть сложным и требовательным к ресурсам. 3. Высокая стоимость: Storm может быть довольно дорогим в эксплуатации, особенно если требуется большой кластер для обработки больших объемов данных. Данный фреймворк может быть использован в различных задачах: от простых преобразований данных до применения алгоритмов машинного обучения.

### **Заключение.**

Таким образом, анализ Big Data в веб-приложениях для настольных игр предоставляет возможность разработчикам получать ценную информацию о поведении пользователей, их предпочтениях и потребностях. Это позволяет улучшать качество игр и увеличивать их популярность среди широкой аудитории. Развитие технологий Big Data открывает новые перспективы для индустрии настольных игр, что делает ее более привлекательной для игроков и инвесторов.

### **Список литературы**

- [1] How to Build Successful Mobile & Web Apps using Big Data? [Электронный ресурс]. - Адрес: <https://www.way2smile.ae/blog/build-successful-mobile-web-apps-using-big-data/>
- [2] Stockfish (chess) [Электронный ресурс]. - Адрес: [https://en.wikipedia.org/wiki/Stockfish\\_\(chess\)](https://en.wikipedia.org/wiki/Stockfish_(chess))
- [3] Game Theory: How Stockfish Mastered Chess [Электронный ресурс]. – Адрес: <https://blogs.cornell.edu/info2040/2022/09/30/game-theory-how-stockfish-mastered-chess/>
- [4] Chess.com [Электронный ресурс]. – Адрес: <https://www.chess.com/>
- [5]. How Many People Play Tabletop Games in 2023? (User & Growth Stats) [Электронный ресурс]. - Адрес: <https://fictionhorizon.com/how-many-people-play-tabletop-games/>
- [6]. Apache Storm: Why use Apache Storm? [Электронный ресурс]. – Адрес: <https://storm.apache.org/>

## **BIG DATA ANALYSIS IN WEB APPS FOR BOARD GAMES**

***G.A. Trofimuk***

*Student of the BSUIR*

***P.A. Arkhireenko***

*Student of the BSUIR*

***A.N. Markov***

*Senior lecturer of the  
department,  
Deputy head of the Center for  
Informatization and Innovative  
Developments*

*Department of Computer Science*

*Faculty of Computer Systems and Networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: gleb.trofimuk@gmail.com*

**Abstract.** Ways of using Big Data analysis in web applications for board games are shown. Examples of ready-made solutions and promising ideas for web applications based on big data analysis related to board games are presented. In addition, a common framework for working with Big Data called Apache Storm is considered for Big Data analysis in web applications.

**Keywords:** Big Data, board games, web apps, Apache Storm.

УДК 004

## ОРГАНИЗАЦИЯ РАЦИОНАЛЬНОЙ СТРУКТУРЫ УПРАВЛЕНИЯ ПЕРЕВОЗОЧНЫМ ПРОЦЕССОМ В ЖЕЛЕЗНОДОРОЖНОМ ТРАНСПОРТЕ



**Ж.Т. Усмонов**

Доцент кафедры Технология почтовой связи ТУИТ имени Мухаммада ал-Хоразмий кандидат технических наук, доцент  
[tax2011@inbox.ru](mailto:tax2011@inbox.ru)



**Т.Б. Джураев**

старший преподаватель кафедры информационных технологии ТУИТ имени Мухаммада ал-Хоразмий,  
[tdjuraev@inbox.ru](mailto:tdjuraev@inbox.ru)



**А.Э. Кувнаков**

Старший преподаватель кафедры информационных технологии ТУИТ имени Мухаммада ал-Хоразмий,  
[avazavaz@umail.uz](mailto:avazavaz@umail.uz)

### **Ж.Т. Усмонов**

Окончил Ташкентского университета информационных технологий. Область научных интересов связана с разработкой Информационных систем, базы данных, мониторинг, менеджмент и оптимизация различных систем

### **Т.Б. Джураев**

Окончил Ташкентского университета информационных технологий. Область научных интересов связана с разработкой Информационных систем, базы данных, мониторинг, менеджмент и оптимизация различных систем

### **А.Э. Кувнаков**

Окончил Ташкентского университета информационных технологий. Область научных интересов связана с разработкой Информационных систем, базы данных, мониторинг, менеджмент и оптимизация различных систем

**Аннотация.** В статье исследуется вопрос оптимизации перевозки грузов в железнодорожном транспорте. Авторы обращают внимание на необходимость организации рациональной структуры управления перевозочным процессом, чтобы обеспечить эффективную работу всей системы. В статье также рассматриваются принципы перевозки грузов в железнодорожном транспорте и проводится анализ ресурсов, необходимых для обеспечения данного процесса.

**Ключевые слова:** структура управления, перевозочный процесс, транспортирование грузов, транспортные операции, формирование грузов.

### **Введение.**

В условиях интенсивного развития информационно-коммуникационных технологий создания новых и совершенствования существующих систем управления сложными организационно-техническими системами, такими как железная дорога, одной из центральных проблем является обоснование и формирование рациональной организационной структуры управления. В этой связи особую важность имеет определение основных принципов построения систем управления процессом формирования поездов.

Важным выводом данного исследования является то, что оптимальную организацию формирования поездов следует рассматривать как одну из подсистем, входящих в большую систему оптимальной организации эксплуатационной работы железнодорожного транспорта. Кроме того, формирование организационной структуры управления

предопределяет «перевозочный процесс» в виде иерархической структуры, в которой неделимой частью является единичная операция перевозки грузов. Основные принципы организации этой структуры должны быть связаны с повышением эффективности и улучшением качества перевозочных услуг.

Таким образом, данное исследование представляет значимый вклад в область организации перевозочного процесса в железнодорожном транспорте и может быть полезным для специалистов, занимающихся вопросами управления этой отраслью транспорта.

#### Основная часть.

Совокупность отдельных зон транспортирования грузов можно рассматривать как участки, а несколько участков - как транспортную железнодорожную магистраль. Перевозочный процесс характеризуется внутренними и внешними связями. Внутренние связи - это связи между отдельными транспортными участками, магистралями.

Внешние связи с различными директивными, снабженческими, плановыми и другими организациями предоставляют транспорту материальные и денежные фонды, кредиты, сведения о новых достижениях науки и техники и пр. Они участвуют в оснащении перевозочного процесса оборудованием, комплектуют его и др. ресурсами. В таблице приведены данные об уровнях управления и управляемых объектах.

Таблица 1. Уровни управления объектами железнодорожного транспорта

Уровни управления	Объекты управления	
	материально-технологический	техничко-экономический
I	Производственный процесс	Производство: а) основное; б) вспомогательное.
II	Оборудование	Основные фонды
III	Материальные и энергетические потоки	Оборотные фонды
IV	Рабочая сила	Трудовые ресурсы
V	Транспортные услуги, готовая продукция	Фонды обращения
VI	Интегральная — подсистема взаимосвязи объектов управления	-

Функционирование перевозочного процесса можно рассматривать как многоцелевой результат, в котором выделяются главные и дополнительные цели, определяемые параметрическими характеристиками  $P_j$  и  $P_j$ .

Совокупность всех целей, определяемая характеристиками  $P_j$  и  $P_j$  представляет собой единую динамическую систему целей ( $\{F_j\}$ ). Перевозочный процесс на ЖТ представляется как множество целей  $\{F\}$ , имеющий иерархическую структуру, где по мере перехода от низшего ранга целей -  $\{F\}_k$  к целям более высокого ранга  $\{F\}_{k+1}$  происходит усложнение по показателям реализации и потребностям в ресурсах для достижения целей ранга, а также уменьшение объема технико-экономической информации (ТЭИ), необходимой для реализации решений задач управления по сравнению с потребностями в ресурсах и её использовании для целей более низких рангов.

Множество целей  $\{F\}$  можно рассматривать в виде совокупности целей по отдельным объектам управления на каждом из уровней иерархии структуры -  $\{F\}_{k-j}$ .

При реализации множества целей  $\{F_{nn}\}$  на различных ступенях управления необходимо выявить перечень задач управления –  $Z_{nn}$ .

Но поскольку множество целей  $\{F_{nn}\}$  характеризуется иерархической структурой с  $K_j$  ранжируемыми ступенями, каждая из которых – множество  $\{F_{nn}\}$  совокупность задач управления –  $Z_{nn}$  представляет собой множество, строение которого подобно множеству  $\{F_{nn}\}$  т. е.  $Z_{nn}$  можно рассматривать в виде иерархической лестницы с количеством рангов, равным количеству рангов –  $\{F_{nn}\}$ .

По мере перехода от низшего ранга задач  $\{Z_{nn}\}_{k-1}$  к более высоким  $\{Z_{nn}\}_k$  происходит их «усложнение» по реализации, что выражается в необходимости разработки совокупности решений  $\{P_{nn}\}$ , позволяющей учитывать все факторы и условия, влияющие на реализацию задач при достижении поставленных целей.

Между целями функционирования задачами управления  $\{Z_{nn}\}$  и решениями по ним  $\{P_{nn}\}$  существует причинно-следственная связь.

Реализация задач управления в первую очередь направлена на восстановления труда, предметов труда и орудий труда. Во-вторых, реализация задач имеет своей целью обеспечение четкого взаимодействия перевозочного процесса внешней средой. Таким образом, множество задач управления  $\{Z_{nn}\}$  включает в себя подмножество задач материально-технического обеспечения  $Z_{nn-мт}$ .

При различных внешних и внутренних возмущениях возможно различное потребление предметов труда и использование орудий труда. По этому для нормального функционирования перевозочного процесса необходимо наличие запаса ресурсов. Однако наличие такого буфера еще не снимает возможности нарушения цикла перевозочного процесса железнодорожного транспорта.

Следовательно, для рационального управления этим процессом необходима реализация условий, которые могли бы предупредить такое нарушение. Это может быть достигнуто путем реализации специальных задач технико-экономического планирования  $Z_{nn-мэт}$ . Решение задач последних позволить ослабить воздействие внешней среды на перевозочный процесс и сохранить его в уравновешенном состоянии.

При реализации перевозочного процесса существует необходимость решения задач совершенствования и прогнозирования  $Z_{nn-с}$ , а также требуется осуществлять контроль за его состоянием, ибо четкого взаимодействия отдельных элементов перевозочного процесса, как показывает практика, можно добиться только при непрерывном анализе фактических и требуемых состояний  $Z_{nn-у}$ . Поэтому в комплексе решаемых задач должны быть задачи оперативного характера  $Z_{nn-оу}$ .

Усложнение связей между отдельными транспортными операциями с внешней средой выдвигает задачи организационного характера  $Z_{nn-о0}$ .

Следует отметить, что организация нормального функционирования перевозочного процесса требует реализации особой группы задач технической подготовки перевозочного процесса  $Z_{nn-мт}$ .

При этом, для оптимальной реализации перевозочного процесса необходимо решать следующие задачи управления  $\{Z_{nn}\}$ : планирования; технической подготовки; материально-технического обеспечения, оперативно - производственного управления; совершенствования производства и прогнозирования; учета и анализа; организационной подготовки.

Рассматриваемая система задач  $\{Z_{nn}\}$  функционально связана с системой решений



$\{P_{nn}\}$ , которая включает в себя множество решений  $\{P_j\}$  задач управления  $Z_{nn}$ . Множество  $P_{nn}$  можно представить как совокупность подмножеств решений отдельных задач управления  $Z_{nn}$ .

Поэтому совокупность решений целесообразно рассматривать в виде иерархии решений  $P_{nn}$ , которая подобна иерархической структуре системы целей  $\{F_{nn}\}$  и задач управления  $\{Z_{nn}\}$ .

В практике задачи управления перевозочным процессом на основе разработки оптимального ПФП являются основными и требуют своего решения в кратчайшие сроки и при наименьших затратах.

Исходя из этого поиск оптимальных решений  $P_{nn}$  связан с выявлением перечня и оценкой ресурсов обеспечения, поскольку качество управления перевозочным процессом зависит от тех ограничений, которые накладываются внешней средой. Одними из них являются условия обеспечения. Количественная оценка ресурсов этих условий представляется нам как одно из исходных данных при поиске решений.

Точность оценки параметров условий обеспечения оказывает определенное влияние на характер принимаемых решений и на процедуру их исполнения. Следует отметить, чем шире перечень рассмотренных и использованных условий, тем достовернее оценка принятых решений, поскольку это обстоятельство позволяет полнее определять направление воздействий условий обеспечения на принятие решения.

Для локальной задачи управления можно предположить практическую неограниченность ресурсов обеспечения в том смысле, что благодаря возможности перераспределения они могут быть получены в требуемых количествах. В то же время, чем выше уровень иерархии, тем больше есть оснований предполагать, что ресурсы будут ограничены и, таким образом, при поиске решений на этих уровнях ограниченность ресурсов обеспечения подлежит постоянному учету, и выработку решений следует определять как функцию от этих условий (1).

В общем виде, влияние условий обеспечения на принимаемые решения по любой задаче  $Z_i$  может быть определено как

$$P_i = f(u_1, u_2, \dots, u_n), \quad (1)$$

где, - характеристика принятого решения по задаче управления;

$u_1, u_2, \dots, u_n$  - виды ресурсов условий обеспечения.

Влияние ресурсов условий обеспечения одного вида может быть выявлено как

$$P_i = f(u_1^1, u_1^2, \dots, u_1^k),$$

где  $u_1^1, u_1^2, \dots, u_1^k$  - числовые значения  $u_1$  вида условий обеспечения.

Основными видами условий обеспечения управленческих решений являются: технико-экономическая информация; средства сбора, обработки и передачи информации; трудовая деятельность работников аппарата управления, занятых обработкой информации и осуществляющих разработку решений; алгоритм выработки решений; нормативная информация, определяющая параметры принимаемых решений.

Опыт показывает, что производственным операциям предшествует разработка управленческих решений. От того, каковы будут эти решения, как они будут влиять на ход перевозочных процессов, зависят качество продукции и размеры транспортных услуг. Следовательно, насущна разработка мероприятий направленных на повышение качества принимаемых решений.

Иными словами, существует такая функциональная связь между качеством продукции (качеством транспортных услуг) и показателями принимаемых решений по технологической подготовке перевозочных процессов:

$$K_n = f(P_n),$$

где  $K_n$  - показатели качества продукции;

$P_n$  - показатели качества решения.

Одним из методов поиска наилучших решений является метод образного сопоставления получаемых решений, основанного на субъективном представлении о характеристике рассматриваемого решения.

Однако этот метод не позволяет сопоставляемые решения оценивать с достаточной точностью. Поэтому целесообразно применение других методов и подходов на основе количественных характеристик рассматриваемых решений. Сопоставление количественных данных о решениях применяется при оценке разработок, направленных на совершенствование перевозочных процессов (например, при определении зон размещения новых транспортных предприятий, методов обработки поступающих грузов и т. п.).

При этом эффективность решений оценивается путем сопоставления себестоимости продукции, транспортных услуг и капитальных затрат при внедрении нового предложения и при базовом варианте

$$\Delta = (C_1 + EK_1) - (C_2 + EK_2), \quad (2)$$

где  $C_1, C_2$  - себестоимость продукции (услуг) при базовом и проектируемом вариантах;

$K_1, K_2$  - капитальные затраты (базовые и проектируемые);

$E$  - нормативный коэффициент экономической эффективности.

Формула (2) может быть использована и при определении относительного уровня рассматриваемых решений. В этом случае относительный уровень предлагаемого решения выявляется как частное от деления второго члена формулы (2) на первый

$$K_3 = \frac{C_2 + EK_2}{C_1 + EK_1} = \frac{Ц_2}{Ц_1}, \quad (3)$$

где  $K_3$  - показатель эффективности (качества) решения;

$Ц_1, Ц_2$  - суммарные капитальные и эксплуатационные затраты при базовом варианте и внедрении нового решения.

Следует отметить, что выражения (2) и (3) в определенном смысле не являются корректными, поскольку при этом не рассматриваются затраты на обработку и использование ТЭИ.

Для управления качеством принимаемых решений при осуществлении перевозочных процессов необходимы такие методы, которые бы учитывали особенности процесса поиска и выбора решения. Выбор решений рассматриваются как процесс преобразования ТЭИ.

При этом порядок и число процедур преобразования информации являются величинами условно постоянными. Изменение порядка и количества процедур преобразования ТЭИ имеет место при внедрении новых методов обработки грузов, которые, как правило, приводят к изменению объемов сопроводительной информации.

Зависимость уровня принимаемых решений от уровня влияющих факторов и условий

представлена на рис. 1.

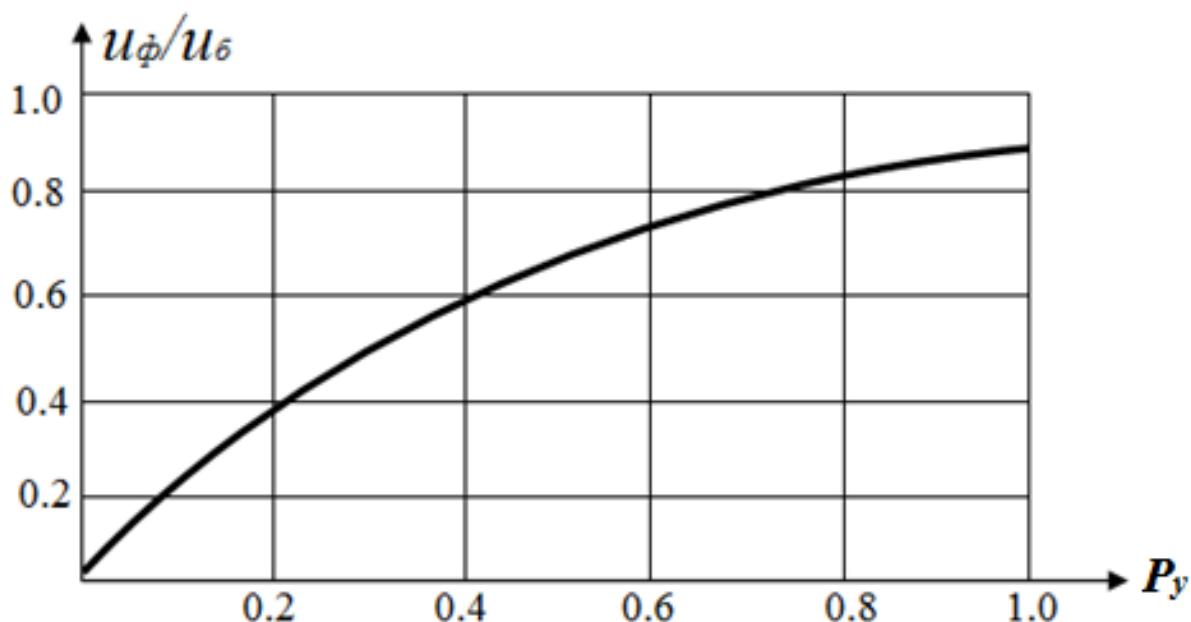


Рисунок 1. Кривая зависимость уровня решений от уровня влияющих факторов и условий

Аналитически зависимость качества управленческих решений от уровня условий обеспечения может быть записана следующим образом:

$$P_y = f(P_1, P_2, \dots, P_j),$$

где  $P_y$  - уровень принимаемых решений;

$P_1, P_2, \dots, P_j$  - условия обеспечения принимаемых решений.

Зная зависимость величин  $P_y$  и  $P_j$  т. е. определив область существования кривых, характеризующих уровень этих условий, можно выявить требуемый уровень условий относительно базовых показателей (рис. 1), определить необходимые затраты на повышение качества ТЭИ, труда работников аппарата управления и др. условий.

Эффективность решений определяется как разность затрат на выработку решений  $Z_p$  и потерь  $P_p$  из-за их несовершенства. Зависимость между затратами  $Z_p$  на разработку и потерями  $P_p$  от реализации решений и их уровнем удобно представить в виде графика (рис. 2.), Последний позволяет также выявить область, где находятся решения, которые следует принимать в ходе управления перевозочными процессами.

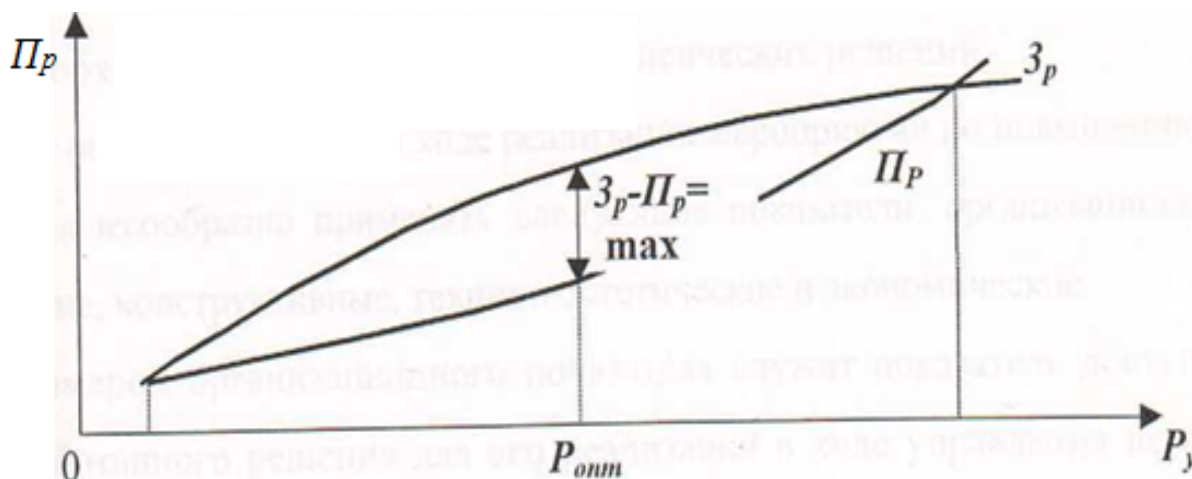


Рисунок 2. График зависимости между уровнем и затратами на выработку решений

Из графика на рис. 2. следует, что максимальной разности величин  $Z_p$  и  $\Pi_p$  соответствует уровень наилучших решений. Определение величин затрат  $Z_p$  и потерь  $\Pi_p$  позволяет заранее знать уровень будущих решений, т. е. прогнозировать уровень решений. Внедрение прогнозирования позволит заранее планировать свойства и параметры новых решений и организовать процесс разработки, оценки и выбора их на производственной основе и снизить влияние субъективных факторов на новые решения.

Предположим, что в начальный момент ввода нормативных документов, уровень рассматриваемого решения  $R_y$  ниже нормативного. Тогда отклонения рассчитанного значения уровня решения от оптимального  $AR$  будет равно величине потерь, вследствие завышения нормативных требований

$$\Delta R_1 = R_{уст} - R_{опт},$$

где  $R_{уст}$  - рассчитанный уровень решения;

$R_{опт}$  - оптимальный уровень решения.

В момент окончания действия нормативных требований значения потерь могут оказаться ниже разработанных оптимальных

$$\Delta R_2 = R_{уст} - R_{опт},$$

В первом случае  $\Delta R_1$ , характеризует величину потерь перевозочного процесса. Для второго случая  $\Delta R_2$  определяется величина дополнительных затрат необходимых при разработке управленческих решений.

#### Заключение.

Формирования организационной структуры управления перевозочным процессом позволяют разрабатывать принципы принятия управленческих решений на всех уровнях перевозочного процесса на основе единых принципов, что имеет существенное значение при автоматизации процессов сбора, обработки и передачи информации при управлении формированием поездов.

### **Список литературы**

- [1] K. Jiyanbekov, J. Usmonov, and S. Azimov, in International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities, ICISCT 2019 (2019)
- [2] J. U. Turdiqulovich, T. D. Boysotovich, and Z. P. Maxmudjonovna, in International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities, ICISCT 2021 (2021)
- [3] J. U. Turdiqulovich, T. B. Djuraev, and Z. M. Pulatova, Journal of Advanced Research in Dynamical and Control Systems **12**, (2020)
- [4] Кувнаков А.Э. Джураев Т.Б. Автомобил йўли орқали юқларни ташиш тизимини автоматлаштириш “Иқтисодиётнинг тармоқларини инновацион ривожланишида ахборот – коммуникация технологияларининг аҳамияти” Республика илмий – техник анжуманининг маърузалар тўплами. Тошкент -2020. –Б. 144-147
- [5] Кувнаков А.Э. Джураев Т.Б. Юлдашев А. Логистика самарадорлиги-нинг рақамли иқтисодиётга таъсири DEUZ) “Рақамли иқтисодиёт: Янги Ўзбекистонни янги технологиялар, платформалар ва бизнес моделлари орқали ривожлантиришнинг янги босқичи” мавзусидаги халқаро илмий – амалий конференция материаллари. 5-қисм. Тошкент -2020. –Б. 63-66
- [6] Djuraev T.B, Alloyofov O.X. “Автомобил транспорти логистикаси фаолиятини Монте-карло усули асосида оптималлаштириш”. “Muhammad al-Xorazmiy avlodlari” ilmiy-amaliy va axborot-tahliliy jurnali. 2(16) /2021, 46-51
- [7] Djuraev T.B., Чориёров Н.К. Модель информационных потоков в диспетчерской службе транспортного процесса. Journal of Integrated Education and Research. Vol. 1 No. 6 (2022): - С. 63-67.

## **ORGANIZING A RATIONAL MANAGEMENT STRUCTURE FOR TRANSPORTATION PROCESSES IN RAILWAY TRANSPORT**

***J.T. Usmonov***

*Associate Professor of the  
Department Postal  
Communication Technology TUIT  
named after Muhammad al-  
Khwarizmi Candidate of  
Technical Sciences, Associate  
Professor*

***T.B. Juraev***

*Senior Lecturer, Department of  
Information Technology, TUIT  
named after Muhammad al-  
Khwarizmi,*

***A.E. Kuvnakov***

*Senior Lecturer, Department  
of Information Technology,  
TUIT named after  
Muhammad al-Khwarizmi,*

*Department of Information and Computer Systems Design  
Faculty of Computer Engineering  
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi,  
Republic of Uzbekistan  
E-mail: tdjurayev@inbox.ru*

**Abstract.** This article explores the question of optimizing freight transportation in railway transport. The authors draw attention to the need for a rational management structure to ensure the efficient operation of the entire system. The article also examines the principles of freight transportation in railway transport and analyzes the resources necessary to support this process.

**Keywords:** management structure, transportation process, transportation of load, transport operations, cargo formation.

УДК [517.98]:

## ПОСТРОЕНИЕ БЛОЧНЫХ РАЗБИЕНИЙ СИСТЕМ БУЛЕВЫХ ФУНКЦИЙ НА ОСНОВЕ ЗАДАЧИ ПОКРЫТИЯ БУЛЕВЫХ МАТРИЦ



**С.Н. Кардаш**  
Старший научный  
сотрудник ОИПИ НАН  
Беларуси, кандидат  
технических наук  
kardash77@gmail.com

**С.Н. Кардаш**

Окончил БГУ им. Ленина. Старший научный сотрудник лаборатории Логического Синтеза ОИПИ НАН Беларуси, к.т.н.

**Аннотация.** Рассматривается задача построения совместных (использующих общие подфункции) разложений систем булевых функций. Используются представления функций в виде дизъюнктивных нормальных форм (ДНФ), полиномов Жегалкина и полиномов Рида-Малера. Предлагается эвристический алгоритм построения разложения минимальной площади. Приводятся результаты экспериментального исследования.

**Ключевые слова:** Дизъюнктивные нормальные формы булевых функций, связанность, полиномы Жегалкина, полиномы Рида-Малера.

### **Введение.**

В настоящее время стремление снизить энергопотребление цифровых систем, реализуемых на элементной базе заказных комплементарных металл-оксид-полупроводниковых схем (КМОП-схем) и систем-на-кристалле стало причиной появления новых и совершенствования известных методов решения задач, связанных с проектированием логических схем.

Синтез логических схем из библиотечных элементов обычно выполняется по оптимизированным двухуровневым либо многоуровневым представлениям систем булевых функций.

Двухуровневыми (И-ИЛИ) представлениями называют представления функций в виде дизъюнктивных нормальных форм (ДНФ), многоуровневыми – различные формы функциональных разложений [1, 2].

Идея использовать связанность (общность) областей определений булевых функций при синтезе многовыходных комбинационных схем предложена в [3].

По существу, выделение связанных функций является одним из приемов логической оптимизации многоуровневых представлений систем функций [4].

Для связанных подсистем функций более эффективно решаются задачи логической оптимизации, например, оптимизации в классе ДНФ [2, 3], оптимизации BDD-представлений [5] и декомпозиции различных видов, например, при построении совместных функциональных разложений.

В работе [6] для многоуровневой оптимизации предложено использовать конъюнктивно-дизъюнктивные разложения матричных форм систем функций, представленных в ДНФ, а также предлагаются алгоритмы построения таких разложений.

Эта идея может быть перенесена на матричные представления функций, заданных в виде полиномов Жегалкина или полиномов Рида-Малера.

Полином Жегалкина называется *многоместная сумма по модулю 2* попарно различных положительных элементарных конъюнкций, т. е. элементарных конъюнкций, не содержащих инверсных литералов.

Полином Жегалкина является канонической формой задания булевой функции – он является единственным для любой булевой функции [7].

В случае, когда полиномы содержат литералы со знаком инверсии (и тогда только с ним), то они носят название полиномов Рида-Малера.

Длиной полинома называется число входящих в него элементарных конъюнкций.

Исходную систему ДНФ, содержащую  $m$  функций,  $n$  аргументов и  $l$  конъюнкций, будем представлять парой матриц – троичной матрицей  $\mathbf{K}$  с размерами  $l \times n$ , нулевые и единичные элементы строк которой представляют инверсные и безынверсные литералы соответствующих конъюнкций, и булевой матрицей  $\mathbf{B}$  с размерами  $l \times m$ , единичные элементы строк которой определяют вхождения этих конъюнкций в соответствующую ДНФ. Площадью такой системы ДНФ будем называть величину  $Q = (m+n) \times l$ .

Введем также в рассмотрение булеву матрицу  $\mathbf{A}$  с размерами  $l \times n$ , элементы которой имеют нулевое значение тогда и только тогда, когда соответствующие им (т.е. находящиеся на тех же позициях) элементы матрицы  $\mathbf{K}$  являются неопределенными.

Аналогичным образом будем представлять и системы полиномов Жегалкина и полиномов Рида-Малера, с той только разницей, что булева матрица  $\mathbf{B}$  будет определять вхождения конъюнкций в соответствующие полиномы.

Под *минором* матрицы будем понимать ее подматрицу, образованную элементами, находящимися на пересечении некоторого подмножества строк и некоторого подмножества столбцов этой матрицы.

Минор, число столбцов которого совпадает с числом столбцов содержащей его матрицы, называется *строчным*, а минор, число строк которого совпадает с числом строк содержащей его матрицы, называется *столбцовым*.

*Вес* минора булевой матрицы, а также веса его строк и столбцов определяются как числа содержащихся в них единичных элементов. Множество миноров булевой матрицы назовем *покрытием* этой матрицы, если каждый ее единичный элемент принадлежит хотя бы одному из этих миноров.

Столбцы и строки булевой матрицы называются *нулевыми* или *ненулевыми* в зависимости от того, все или не все их элементы имеют нулевое значение.

Минор вышеупомянутой матрицы  $\mathbf{B}$  называется *(R,T,S)-ограниченным*, если он содержит не более  $R$  строк и не более  $T$  столбцов, а соответствующий ему (т.е. образованный строками с такими же номерами) строчный минор матрицы  $\mathbf{A}$  имеет не более  $S$  ненулевых столбцов.

Систему ДНФ будем называть *(R,T,S)-связанной*, если числа ее конъюнкций, функций и аргументов не превышают заданных величин  $R$ ,  $T$  и  $S$  соответственно.

Под системой *индексированных* ДНФ будем понимать систему таких ДНФ, которым взаимно однозначно сопоставлены некоторые натуральные числа, называемые *индексами* ДНФ.

Систему индексированных ДНФ будем называть *подсистемой* другой системы индексированных ДНФ, если множества аргументов, конъюнкций и индексов первой являются подмножествами аргументов, конъюнкций и индексов другой системы.

Пусть  $D$  – исходная система индексированных ДНФ,  $\{1, 2, 3, \dots, m\}$  – множество индексов ДНФ этой системы. Множество  $W$  подсистем системы  $D$  назовем ее *дизъюнктивным разложением*, если каждая ДНФ системы  $D$  равняется дизъюнкции всех ДНФ с таким же индексом, имеющих в подсистемах множества  $W$ .

Аналогичным образом определяются соответствующие разложения для систем полиномов Жегалкина и полиномов Рида-Малера.

В дальнейшем будем рассматривать только случай системы ДНФ, имея в виду, что при схемной реализации разложений полиномов Жегалкина и полиномов Рида-Малера вместо операции «дизъюнкция» используется операция «сумма по модулю 2».

**Постановка задачи.** Рассматриваемая задача заключается в построении дизъюнктивного разложения системы  $D$  на  $(R, T, S)$ -связанные подсистемы, где  $R$ ,  $T$  и  $S$  – заданные натуральные числа, с минимальной суммарной площадью.

Эту задачу предлагается свести к нахождению покрытия матрицы  $B$   $(R, T, S)$ -ограниченными минорами с минимальным суммарным весом. Заметим, что эта задача разрешима лишь в том случае, когда *ранг* (число литералов) каждой конъюнкции в системе не превышает  $S$ .

**Описание алгоритма.** Для решения сформулированной задачи предлагается циклический алгоритм, на  $i$ -м шаге которого в матрице  $B$  выбирается  $(R, T, S)$ -ограниченный минор  $B'_i$ , содержащий лишь ненулевые строки, после чего подсистема  $D_i$  системы  $D$ , определяемая этим минором и соответствующим ему строчным минором  $K_i$  матрицы  $K$ , заносится в формируемое решение, а затем единичным элементам матрицы  $B$ , принадлежащим минору  $B'_i$ , присваиваются нулевые значения.

Процесс формирования дизъюнктивного разложения заканчивается, когда в матрице  $B$  не останется единичных элементов.

Минор  $B_i$  булевой матрицы  $B$  будем называть *ненулевым*, если он содержит хотя бы один единичный элемент, а минор этой же матрицы, получаемый в результате удаления из  $B_i$  нулевых столбцов и нулевых строк, назовем *остовом* минора  $B_i$ .

Пусть  $a_j$  –  $j$ -я строка матрицы  $A$ ;  $A_i$  – некоторый строчный минор этой матрицы, не содержащий строки  $a_j$ ;  $A_i^j$  – строчный минор матрицы  $A$ , являющийся результатом включения строки  $a_j$  в минор  $A_i$ , а  $B_i$ ,  $B'_i$  и  $b_j$  – соответствующие минорам  $A_i$ ,  $A_i^j$  и строке  $a_j$  строчные миноры и строка матрицы  $B$ . Строку  $a_j$  будем называть  $S$ -совместимой с минором  $A_i$ , если строка  $b_j$  является ненулевой, а число ненулевых столбцов минора  $A_i^j$  не превышает  $S$ . **Базой** минора  $A_i$  называется булев вектор длины  $n$ , помечающий своими единичными компонентами ненулевые столбцы этого минора. Через  $a_{ij}$  будем обозначать число пар одноименных элементов строки  $a_j$  и базы минора  $A_i$ , имеющих различные значения, а через  $b_{ij}$  – разность между суммой весов  $T$  наиболее "тяжелых" (большого веса) столбцов минора  $A_i^j$  и суммой весов  $T$  наиболее "тяжелых" столбцов минора  $B_i$ .

Формирование минора  $B'_i$  начинается с выделения в матрице  $B$  минора  $B_i$ , образованного той из ее ненулевых строк, которой соответствует наиболее "тяжелая" строка матрицы  $A$ , и выделения в матрице  $A$  соответствующего ему строчного минора  $A_i$ .

Далее осуществляется итерационный процесс, каждая итерация которого заключается в выполнении следующих действий. В начале производятся построчные наращивания минора  $A_i$   $S$ -совместимыми с ним строками матрицы  $A$  и соответствующие им построчные наращивания минора  $B_i$ , причем критерием выбора номера  $j$  очередных включаемых в миноры строк является максимум величины  $(b_{ij} + 1) / (a_{ij} + 1)$ . Эти наращивания заканчиваются, когда либо в матрице  $A$  не окажется  $S$ -совместимых с минором  $A_i$  строк, либо число его строк станет равным  $R$ .

Затем в матрице  $B$  выделяется минор  $B'_i$ , образованный  $T$  наиболее "тяжелыми" столбцами минора  $B_i$ , после чего из  $B'_i$  удаляются нулевые строки, а соответствующие им строки матриц  $A$  и  $B$  удаляются из миноров  $A_i$  и  $B_i$ . Описанный процесс выполняется до тех



пор, пока вес вновь построенного минора  $B'_i$  не окажется равным весу одноименного минора, сформированного на предыдущей итерации, после чего из этих двух миноров в качестве результата выбирается тот, которому соответствует строчный минор матрицы  $A$  большего веса.

#### Экспериментальное исследование.

Предложенный алгоритм реализован на языке программирования C++ с использованием библиотеки классов «булев вектор» и «троичный вектор». Для проверки эффективности предложенного алгоритма был проведен вычислительный эксперимент.

Примеры матричных SF-описаний систем полностью определенных булевых функций были взяты из набора промышленных тестовых примеров, входящих в библиотеку примеров *Berkeley PLA TestSet* [8]. Для каждого них с помощью программ, входящих в систему FLC логической оптимизации [9], были построены матричные представления в виде полиномов Жегалкина и полиномов Рида-Малера.

В ходе эксперимента исследовалось применение предложенного алгоритма для трех форм представления матричных описаний – системы ДНФ (SDF), систем полиномов Жегалкина (PG) и систем полиномов Рида-Малера (PRM). Оценивалась суммарная площадь получаемых разложений.

Результаты эксперимента представлены в таблице, где  $n$  – число переменных,  $m$  – число функций,  $k$  – число элементарных конъюнкций исходной системы ДНФ булевых функций. Через  $Q_{ISX}$  обозначена площадь исходного матричного представления систем булевых функций, через  $Q_{PG}$  обозначена площадь для полиномов Жегалкина. Жирным шрифтом выделены лучшие решения.

В ходе экспериментов для каждого примера осуществлялся автоматический подбор параметра  $T$  (числа функций в блоках разложения) при фиксированных параметрах  $R = k$ ,  $S = n$ . При этом перебирались последовательные значения  $T = 1, 2, 3, \dots, n$  и каждый раз оценивалась число  $N$  блоков и площадь полученного разложения. Процесс останавливался, когда на очередном шаге площадь полученного разложения возрастала. В этот момент и фиксировались параметры  $T$  и  $N$ .

Таблица 1. Результаты работы программы построения функциональных разложений

Имя	n	m	k	Q <sub>ISX</sub>	SDF			Q <sub>ISX</sub>	PG			Q <sub>ISX</sub>	PRM		
					T	N	Q <sub>SDF</sub>		T	N	Q <sub>PG</sub>		T	N	Q <sub>PRM</sub>
Mr2d	14	14	123	3444	4	4	1855	9 884	1	14	5 768	2 660	3	5	<b>1655</b>
TIAL	14	8	640	14080	1	8	<b>8843</b>	96 932	4	2	82 926	81 026	4	8	74488
IN2	19	10	137	<b>3 973</b>	10	1	3973	183 135	9	2	108 630	10 295	4	1	9942
IN0	15	11	138	3 588	11	1	<b>3510</b>	177 190	10	1	144 140	9 074	11	1	9074
ADD6	12	7	1092	20748	1	7	13480	2 508	1	7	1 572	2 508	1	7	<b>1572</b>
B2	16	17	110	3630	17	1	<b>3630</b>	101 970	17	1	101 970	35 178	17	1	35178
T3	12	8	152	3040	1	1	1845	18 140	4	2	14 798	1 020	8	1	<b>1020</b>
RADD	8	5	120	1560	5	1	1008	442	5	1	<b>272</b>	546	5	1	349

#### Заключение.

Уменьшить площадь получаемых в процессе синтеза комбинационных нерегулярных логических схем можно с помощью логической оптимизации исходных

описаний. Хорошим средством для улучшения конечных схемных решений может служить предварительная обработка исходных систем с помощью программ построения функциональных разложений. Предложенный в работе алгоритм построения таких разложений является быстродействующим и позволяет в ряде случаев получать более эффективные схемные решения. Результаты экспериментального исследования показали, что использование разработанных программных средств может сократить площадь получаемых решений до полутора раз.

#### **Список литературы**

- [1] Закревский, А. Д. Логические основы проектирования дискретных устройств / А. Д. Закревский, Ю. В. Потгосин, Л. Д. Черемисинова. – М.: Физматлит, 2007. – 592 с.
- [2] Бибило, П. Н. Декомпозиция булевых функций на основе решения логических уравнений / П. Н. Бибило. – Минск: Беларус. навука, 2009. – 211 с.
- [3] Кузнецов, О. П. О программной реализации логических функций и автоматов / О. П. Кузнецов // Автоматика и телемеханика. – 1977. – № 7. – С. 63–74.
- [4] Бибило, П.Н. Дизъюнктивно-конъюнктивные разложения систем полностью определенных булевых функций / П.Н. Бибило, С.Н. Кардаш / Доклады Восьмой Международной научной конференции «Танаевские чтения», 27–30 марта 2018 г. – Минск, ОИПИ НАН Беларуси, 2018 г. – С. 28–32.
- [5] Akers, S. V. Binary decision diagrams / S. V. Akers // IEEE Trans. on Computers. – 1978. – Vol. C-27, no. 6. – P. 509–516.
- [6] Кардаш, С. Н. Функциональные разложения систем булевых функций / С. Н. Кардаш // VIII Международная научно-практическая конференция «BIG DATA and Advanced Analytics» (BIG DATA 2022): Материалы междунар. науч. конф., (Республика Беларусь, Минск, 11-12 мая 2022 г.). – Минск: БГУИР, 2022. – С. 142–150.
- [7] Закревский, А. Д., Торопов Н.Р. Полиномиальная реализация частичных булевых функций. – Минск: Минск: Ин-т техн. Кибернетики НАН Беларуси, 2001. – 200 с.
- [8] Berkeley PLA test set [Electronic resource]. Mode of access: <http://www1.cs.columbia.edu/~cs6861/sis/espresso-examples/>. Date of access: 9.12.2015.
- [9] Бибило П.Н., Романов В.И. Логическое проектирование дискретных устройств с использованием продукционно-фреймовой модели представления знаний. Изд. 2-е, испр. – М.: ЛЕНАНД, 2014, 256 с. Бибило, П.Н. Кремниевая компиляция заказных СБИС. - Минск: Институт технической кибернетики АН Беларуси, 1996. - 268 с.

## **CONSTRUCTION OF BLOCK PARTITIONS OF SYSTEMS OF BOOLEAN FUNCTIONS BASED ON THE PROBLEM OF COVERING BOOLEAN MATRIXES**

***S.N.Kardash***

*Senior Research Fellow of the United  
Institute of Informatics Problems of the  
National Academy of Sciences of  
Belarus, PhD*

*Research Institute for Applied Problems of Mathematics and Informatics of the Belarusian State University, Republic of Belarus  
E-mail: kardash77@gmail.com*

**Abstract.** The problem of constructing joint (using common subfunctions) expansions of systems of Boolean functions is considered. Representations of functions in the form of disjunctive normal forms (DNF), Zhegalkin polynomials and Reed-Mahler polynomials are used. Algorithms for constructing the decomposition of the minimum area are proposed. The results of an experimental study are presented.

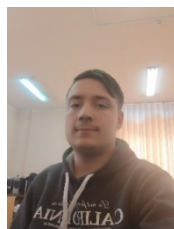
**Keywords:** Disjunctive normal forms of Boolean functions, connectedness, polynomias Zhegalkin, polynomias of Reed-Maler.

УДК 004.021:004.75

## АНАЛИЗ BIG DATA В GAMEMOBILE



**Г.А. Трофимук**  
Студент БГУИР  
gleb.trofimuk@gmail.com



**П.А. Архиреев**  
Студент БГУИР  
arhir.13.07.2002@gmail.com



**А.Н. Марков**  
Старший преподаватель,  
магистр технических наук,  
заместитель начальника Центра  
информатизации и  
инновационных разработок  
БГУИР  
a.n.markov@bsuir.by

**П.А. Архиреев**  
Студент кафедры информатики БГУИР.

**Г.А. Трофимук**  
Студент кафедры информатики БГУИР.

**А.Н. Марков**  
Магистр технических наук, старший преподаватель кафедры ПИКС, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Данная статья исследует причины возникновения больших данных в мобильных приложениях, их особенности и характеристики. Кроме того, для анализа больших данных мобильных приложений авторы рассматривают распространенный фреймворк Spark, который предлагает эффективный и экономичный способ хранения данных в больших объемах, включая слабоструктурированные и неструктурированные данные.

**Ключевые слова:** большие данные, Big Data, Hadoop, Spark, мобильные приложения.

### **Введение.**

Современные информационные системы ориентированы на обработку огромных объемов данных, которые генерируются ежедневно [1].

В научных, промышленных и бизнес-отраслях необходимо уметь эффективно использовать инструменты для обработки доступных Big Data, чтобы получить ценные результаты.

Одним из наиболее активно развивающихся направлений в области информационных технологий, связанных с большими данными, являются мобильные технологии, которые объединяют Интернет вещей, ориентированный на устройства, и мобильные приложения, ориентированные на пользователя [2].

### **Актуальность работы с BIG DATA.**

Big Data - это набор подходов, методов и инструментов для обработки значительных объемов данных с различной степенью структурированности.

Основная цель обработки Big Data заключается в оперативном и эффективном использовании всей доступной информации, которая постоянно растет и меняется в огромных объемах.

Обработка больших данных позволяет работать с объемом информации, который невозможно обработать стандартными аппаратными средствами и инструментами, такими как системы управления реляционными базами данных (БД) [3].

Основные задачи Big Data заключаются в хранении и обработке больших объемов информации, что требует использования специальных технологий и инструментов, предназначенных для работы со сверхбольшим объемом данных.

Со стороны разработчика, Big Data в мобильных играх - крайне важный инструмент для понимания поведения и предпочтений пользователей.

С помощью такого анализа разработчики могут получить ценную информацию о том, как игроки взаимодействуют с игрой, какие элементы им нравятся больше всего, какие сложности возникают и как улучшить игровой процесс.

Основные характеристики Big Data, собираемых в мобильных приложениях, включают в себя [4]:

1. Объем данных, который превышает возможности типичных БД, что требует новых подходов и инструментов для сбора, хранения, обработки и анализа данных.

2. Скорость обработки данных в режиме реального времени, что позволяет быстро реагировать на текущую информацию, ограниченную временным периодом приложения.

3. Многообразие источников данных, которые содержат различные типы и форматы данных, включая структурированные, полуструктурированные и неструктурированные данные, извлекаемые из умных устройств, сенсоров, мобильных приложений и социальных сетей. Это позволяет интегрировать большое количество источников данных для обработки и анализа.

4. Недостоверные данные могут затруднить их анализ, поэтому достоверность информации является одним из наиболее важных критериев для пользователей.

5. Большие данные имеют ценность благодаря своей способности улучшать процессы составления отчетности, организации бизнес-процессов, оптимизации расходов компаний и других сфер деятельности, где накопление информации играет важную роль.

Из-за большого количества разнородной информации, Big Data собранные в результате работы мобильного приложения, могут содержать как множество полезной информации, так и наоборот – мусора.

Без правильного анализа она не может быть отфильтрована в полной мере, ведь грамотный анализ больших данных в данной отрасли может помочь разработчикам создавать более привлекательные игры, лучше соответствующие потребностям пользователей и способствующие удержанию игроков в игре на больший период времени.

Именно массивное развитие мобильных устройств и приложений привело к возникновению больших объемов данных, которые сейчас пользователи получают в огромных количествах.

Для обработки такого объема данных, разработчикам мобильных приложений необходимо создавать сложные системы хранения, использующие различные модели данных [5].

В этом случае, для разработки приложений, обрабатывающих большие объемы данных, программисты прибегают к применению соответствующих архитектурных шаблонов проектирования.

Поскольку данные являются ценным активом, необходимо обеспечить надежную систему защиты, которая должна учитываться при любом выборе архитектуры. Кроме того, при разработке систем, предназначенных для системного анализа больших объемов данных в мобильных приложениях, необходимо учитывать ряд факторов [6]:

1. Для успешного анализа данных необходимо предварительно обработать их, поскольку исходная информация неупорядочена и требует очистки, объединения, изменения, перемешивания и других манипуляций для извлечения полезной информации.

2. В сфере обработки больших данных ключевым понятием является итерация.

Для проведения анализа и моделирования часто требуются многократные проходы по одной и той же информации в связи со статистическими процедурами и алгоритмами машинного

обучения. Однако, если при каждом обращении к диску фреймворку необходимо читать один и тот же набор сведений, это может привести к задержкам, которые замедляют всю работу.

3. Для полного использования возможностей Big Data необходимо преобразование хорошо функционирующей модели из регрессионных коэффициентов, которые хранятся в текстовом файле на устройстве разработчика, в информационные приложения, где модели становятся частью сервиса, работающего в эксплуатационном режиме и требующего периодической перестройки.

Конечно, при этом так же следует учитывать вопросы конфиденциальности данных и обеспечивать защищенность личной информации самих игроков, так как нет в мире системы, что невозможно взломать.

Однако, если выполнять все необходимые требования по безопасности, то анализ Big Data может стать мощным инструментом для улучшения мобильных игр и повышения качества пользовательского опыта.

Необходимо отметить, что на данное время уже есть отличные примеры грамотной реализации этой технологии:

- Компания King Digital Entertainment, известная своей игрой "Candy Crush Saga", использует анализ данных для мониторинга поведения своих пользователей.

Они собирают данные о том, как часто игроки проходят уровни и покупают игровые предметы, чтобы определить, насколько привлекательной является игра для пользователей.

Эта информация помогает компании разработать более интересные игры и удерживать пользователей в своих приложениях.

- Playtika - крупнейший разработчик социальных казино-игр. Компания использует анализ данных для определения того, какие типы игр работают лучше всего для различных групп пользователей. Они также анализируют данные о том, как игроки взаимодействуют с играми, чтобы улучшить пользовательский опыт и повысить удовлетворенность пользователями.

- Machine Zone - разработчик мобильных стратегических игр, использует анализ Big Data для создания целенаправленной рекламы и улучшения процесса монетизации.

Они используют данные о том, как пользователи играют и покупают игровые предметы, чтобы настроить рекламу и предложения, которые лучше соответствуют интересам и потребностям пользователей.

Эти примеры демонстрируют, что при умелом подходе к Big Data в мобильных играх, полученные после анализа данные - это не только хороший инструмент заработка, но и возможность совершенствовать свой продукт, опираясь на данные, которые предоставляют ваши же клиенты.

### **Принципы функционирования Apache Spark.**

При разработке мобильных приложений, которые предназначены для практического использования, разработчики занимаются операционным анализом и превращают свои модели в сервисы, которые могут быть использованы для решения различных задач.

Такие приложения постоянно отслеживаются на производительность, контроль SLA и время доступности сервиса, а также совершенствуются различные характеристики.

А для этого необходим фреймворк, который обеспечивает удобство моделирования и годится для разработки промышленных систем разного типа. Apache Spark является одним из таких фреймворков.

Он обладает открытым исходным кодом и сочетает в себе элегантную модель написания кода поверх кластера машин с механизмом распределения программ. Он так же является частью экосистемы проектов Hadoop [7].

Apache Spark представляет собой распределенное кластерное приложение, которое используется для обработки данных. Оно также является универсальной платформой, которая поддерживает несколько режимов обработки данных, включая пакетную и потоковую обработку, а также обработку графов и SQL-запросы [1].

Кроме того, Spark предоставляет различные инструменты для этих целей. Эта платформа специально разработана для создания распределенных приложений, которые могут обрабатывать данные. В связи с этим она предлагает фиксированную master-slave архитектуру, основными компонентами которой являются:

1. Driver, который отвечает за управление вычислениями, в том числе распределение задач на вычислительные ресурсы с помощью executor'ов, учитывая информацию об их использовании. Он также генерирует задачи и перезапускает их в случае отказов, контролирует состояние обработки и управляет вычислительными ресурсами, выделяя и возвращая их менеджеру ресурсов.

2. Executor - компонент, который отвечает за выполнение вычислений на ресурсах того вычислительного узла, где располагается информация.

Он обеспечивает хранение информации и предоставление доступа к ней для выполнения обработки, включая десериализацию или загрузку данных с диска.

Он также контролирует потребление вычислительных ресурсов и запускает задачи, фиксируя их результаты.

Spark, подобно своему предшественнику MapReduce, обеспечивает отказоустойчивость и линейную масштабируемость, но расширяет их возможности в нескольких важных направлениях [6]:

- Выполнение универсального ориентированного ациклического графа, что позволяет передавать промежуточные результаты непосредственно на следующий этап конвейера.

- Широкий спектр преобразований, которые позволяют пользователям задавать вычисления естественным образом. Spark имеет потоковый API, который может определять конвейеры всего в нескольких строках кода.

- Обработка в оперативной памяти, что позволяет разработчикам взаимодействовать с любым местом в обрабатываемом конвейере в машинной памяти кластера. Это избавляет от необходимости повторного вычисления данных или чтения их с диска на следующих этапах и расширяет возможности использования механизмов для распределенной обработки.

- Spark - оптимальный выбор для высокоитеративных алгоритмов и приложений, которые быстро реагируют на запросы пользователей благодаря просмотру больших наборов данных в памяти. Он также характеризуется отличной интеграцией со многими инструментами из экосистемы Hadoop.

- Исходя из представленных данных, можно подытожить несколько преимуществ, что позволяют нам остановиться на выборе именно этого фреймворка для работы с Big Data, получаемых из мобильных приложений:

- Высокая скорость обработки данных: Apache Spark может обрабатывать данные в режиме реального времени, что позволяет быстро анализировать большие объемы данных в мобильных приложениях.

- Универсальность: Apache Spark поддерживает множество языков программирования, включая Java, Scala, Python и R, что делает его универсальным и легко интегрируемым в различные проекты.

- Распределенная обработка данных: Apache Spark использует распределенную обработку данных, что позволяет обрабатывать большие объемы данных на кластерах из множества узлов.

- Масштабируемость: Apache Spark легко масштабируется, что позволяет увеличивать объемы обрабатываемых данных и количество узлов в кластере.

- Разнообразие возможностей: Apache Spark предоставляет широкий набор библиотек и инструментов для анализа данных, включая машинное обучение, графовые алгоритмы, обработку потоков данных и многое другое.

- Открытый исходный код: Apache Spark является проектом с открытым исходным кодом, что делает его доступным и гибким для использования в различных проектах.

### **Заключение.**

Работа с большими объемами данных в мобильных приложениях представляет ряд сложностей, таких как недостаточная инфраструктурная мощность, необходимость обеспечения должного уровня безопасности и организационные трудности при внедрении инновационных методов сбора данных из-за ограниченных ресурсов.

Применение фреймворка Apache Spark позволяет частично решить данные проблемы, улучшив отказоустойчивость и скорость обработки данных, а также эффективно выполнять итеративные алгоритмы благодаря кэшированию результатов в памяти и расширения технологий экосистемы Hadoop.

### **Список литературы**

- [1] Бутаков Н.А., Петров М.В., Насонов Д. Обработка больших данных с Apache Spark. – СПб.: Университет ИТМО, 2019. – 50 с.
- [2] Комаров М., Кучерявый Е., Источники Больших Данных и современные способы хранения данных // Тезисы докладов конференции «Большие Данные в национальной экономике» / под ред. Дубовой Н.А. – М.: «Открытые системы», 2014. – С. 9-11.
- [3] Серебряков М.А. Исследование архитектур построения и алгоритмов работы систем хранения и анализа больших данных: выпуск. квалиф. работа: 09.04.03 / М.А. Серебряков; Тольятт. гос. ун-т. – Тольятти, 2019. – 97 с.
- [4] Кобзаренко Д.Н., Мустафаев А.Г. Учебное пособие дисциплины «Анализ больших данных» для направления подготовки 38.03.05 «Бизнес-информатика», профиль «Электронный бизнес». – Махачкала: ДГУНХ, 2019. – 107 с.
- [5] Бойкова М.А., Петрова С.Ю. Пример реализации интероперабельности больших данных для мобильных приложений // Вестник Новгородского государственного университета. – 2017. – № 6 (104). – С. 6-10.
- [6] Риза С., Лезерсон У., Оуэн Ш., Уиллис Д. Spark для профессионалов: современные паттерны обработки больших данных. – М.: Из-во Питер, 2017. – 272 с.
- [7] Стаселько И.Д., Сычев А.Ю., Протасов А.П., Позняков Т.Д., Алексеев Ю.И. Форматы файлов больших данных: хранение данных в экосистеме Hadoop // Big Data and Advanced Analytics. – 2020. – № 6-2. – С. 335-339.

## **BIG DATA ANALYSIS IN GAMEMOBILE**

**G.A. Trofimuk**

*Student of the BSUIR*

**P.A. Arkhireenko**

*Student of the BSUIR*

**A.N. Markov**

*Senior lecturer of the department,  
Deputy head of the Center for  
Informatization and Innovative  
Developments*

*Department of Computer Science*

*Faculty of Computer Systems and Networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: arhir.13.07.2002@gmail.com*

**Abstract.** This article explores the reasons for the emergence of big data in mobile applications, their peculiarities, and characteristics. In addition, to analyze big data in mobile applications, the authors consider the widely used Spark framework, which offers an efficient and cost-effective way to store large volumes of data, including weakly structured and unstructured data.

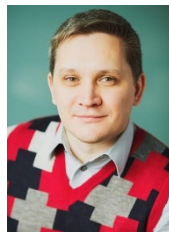
**Keywords:** Big Data, Hadoop, Spark, Mobile applications.

УДК 625

## СТАТИСТИЧЕСКИЙ АНАЛИЗ МЕЖКОМПОНЕНТНЫХ ФАЗОВЫХ ОТНОШЕНИЙ



**Д.А. Кечик**  
Ассистент кафедры  
информационных  
радиотехнологий БГУИР,  
магистр технических наук  
[ya.dan.kechik@yandex.by](mailto:ya.dan.kechik@yandex.by)



**И.Г. Давыдов**  
Доцент кафедры  
информационных  
радиотехнологий БГУИР,  
кандидат технических наук,  
доцент  
[davydov\\_ig@bsuir.by](mailto:davydov_ig@bsuir.by)

### **Д.А. Кечик**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с построением алгоритмов обработки сигналов вибрации с целью диагностики оборудования.

### **И.Г. Давыдов**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с построением алгоритмов обработки сигналов вибрации с целью диагностики оборудования.

**Аннотация.** В настоящей работе представлена аппроксимация распределения межкомпонентных фазовых отношений (МКФО) составляющих полигармонического сигнала, зашумлённого гауссовским шумом. Показано, что для высоких отношений сигнал-шум закон распределения МКФО близок к нормальному.

**Ключевые слова:** межкомпонентные фазовые отношения, цифровая обработка сигналов, статистика, тест Колмогорова-Смирнова.

### **Введение.**

Неразрушающий контроль технического состояния промышленного оборудования существенно повышает эффективность его использования за счёт удлинения межремонтного интервала [1]. Одним из методов неразрушающего контроля является анализ сигналов вибрации агрегата.

Одной из распространённых проблем сборки оборудования, приводящая к значительному росту частоты отказа его узлов, является расцентровка соединяемых валов [2]. Ряд дефектов, в том числе расцентровка валов, характеризуется наличием в спектре виброускорения трёх гармонических составляющих с частотами  $F_0$ ,  $2F_0$ ,  $3F_0$ , где  $F_0$  – частота вращения вала. Помимо обнаружения расцентровки валов, остро стоит проблема определения её величины.

Для различения вида и степени расцентровки валов, в работе [3] предложено использовать межкомпонентные фазовые отношения (МКФО) узкополосных составляющих вибрации. МКФО представляют собой линейные комбинации фаз составляющих. Проанализировано распределение величин, вычисленных как усреднённые МКФО отфильтрованных компонентов вибрации. Эмпирическая плотность распределения этих величин близка к гауссовой, а их математическое ожидание зависит от величины смещения валов.



Цель данной работы – получить теоретическое распределение значений МКФО аргументов Фурье-коэффициентов, соответствующих гармоническим составляющим сигнала; построить зависимость среднеквадратического отклонения (СКО) МКФО от отношения сигнал-шум (ОСШ) и длины реализации. Это поможет в дальнейшем построить алгоритм принятия решения о виде и степени дефекта, оценить его параметры.

### Статистические распределения фазы.

Представим вибрационный сигнал диагностируемого оборудования в виде:

$$x(t) = \sum_{h=1}^H A(h) \cos(2\pi h F_0 t + \phi(h)) + n(t) = \sum_{h=1}^H A(h) \cos(2\pi h F_0 t + \hat{\phi}(h)), \quad (1)$$

$$\phi(h) = h\phi_0 + \theta(h) = E[\hat{\phi}(h)]$$

где  $H$  – число квазигармонических колебаний,  $A(h)$  – амплитуда  $h$ -го колебания,  $\phi_0$  – начальная постоянная фаза компонентов вибрации, обусловленная случайным моментом начала записи. Значения  $A(h)$  и  $\theta(h)$  считаются априори известными константами, как и дисперсия шума  $\sigma^2$ . Амплитуды составляющих сигнала вибрации при различных величинах расцентровки валов могут быть получены решением системы дифференциальных уравнений [4] или получены экспериментально [5]. Зависимость величины  $\theta(h)$  от величины расцентровки была установлена ранее экспериментально [3].

Рассмотрим распределение спектральных отсчётов  $X(hF_0)$ ,  $h = 1, 2, \dots, H$ :

$$X(f) = \sum_{t=0}^{T-1} x(t) \cdot \exp(-2\pi j f \cdot t). \quad (2)$$

В силу линейности преобразования Фурье (ПФ),  $X(hF_0) = A(h) + N(hF_0)$ , где  $N(hF_0)$  – ПФ шума:

$$N(f) = \sum_{t=0}^{T-1} n(t) \cdot \exp(-2\pi j f \cdot t). \quad (3)$$

Согласно [6], мнимые и действительные части спектральных отсчётов  $N(f)$  при условии стационарности  $n(t)$  распределены асимптотически по нормальному закону:

$$D(\Re[N(f)]) = D(\Im[N(f)]) = N [0, T \cdot P(f)] \quad (4)$$

где  $T$  – дискретное время,  $P(f)$  – спектральная плотность мощности, определяемая как:

$$P(f) = \sum_{\tau=-\infty}^{\infty} \exp(2\pi j f \tau) \cdot c(\tau) \quad (5)$$

$$c(\tau) = \text{cov}[X(t+\tau), X(t)], \text{cov}(X, Y) = E[X - E(X)] \overline{E[Y - E(Y)]} \quad (6)$$

Т.к.  $n(t)$  – белый гауссов шум, то  $c(\tau) = \sigma^2 \delta(0)$ , соответственно,  $P(f) = P_0 = \sigma^2 / F_s$ , где  $\sigma^2$  – мощность шума,  $F_s$  – частота дискретизации, тогда:

$$D(\Re[N(f)]) = D(\Im[N(f)]) = N [0, T \cdot P_0] \quad (7)$$

Если  $X(f)$  нормировать по числу отсчётов  $T$ , дисперсия Фурье-отсчётов уменьшится в  $T^2$  раз:

$$D(\Re[N(f)]) = D(\Im[N(f)]) = N[0, \sigma_s^2], \sigma_s^2 = T^{-1} \cdot P_0 \quad (8)$$

Распределение аргумента комплексной величины  $X(hF_0)$  рассчитывается через совместное распределение мнимой и действительной частей [7]:

$$D(\Re[X(hF_0)], \Im[X(hF_0)]) = (2\pi\sigma_s^2)^{-1} \cdot \exp[-(\Re[X(hF_0)] - a)^2 + (\Im[X(hF_0)] - b)^2] / (2\sigma_s), a = A(h)\cos(\phi(h)), b = A(h)\sin(\phi(h)) \quad (9)$$

которое после преобразования декартовых координат в полярные с учётом

$$\Re[X(hF_0)] = r(h)\cos\hat{\phi}(h), \quad (10)$$

$$\Im[X(hF_0)] = r(h)\sin\hat{\phi}(h),$$

где  $r(h)$ ,  $\hat{\phi}(h)$  – амплитуда и фаза Фурье-отсчёта  $h$ -й гармоники соответственно, преобразуется в совместное распределение амплитуды и фазы:

$$D(r(h), \hat{\phi}(h)) = r(2\pi\sigma_s^2)^{-1} \cdot \exp[-(r(h)\cos\hat{\phi}(h))^2 + (r(h)\sin\hat{\phi}(h) - b)^2] / (2\sigma_s) \quad (11)$$

откуда распределение фазы получено интегрированием (11) по амплитуде:

$$D(\hat{\phi}(h)) = \int_0^\infty D(r(h), \hat{\phi}(h)) dr = \frac{\exp(-A^2(h)/2\sigma_s^2)}{2\pi} + \frac{A\cos u}{2\sigma_s\sqrt{2\pi}} \cdot \left[ 1 + \operatorname{erf}\left(\frac{A\cos u}{\sqrt{2\sigma_s}}\right) \right] \cdot \exp\left(-\frac{A^2}{2\sigma_s^2} \sin^2 u\right), u = \hat{\phi}(h) - \phi(h) \quad (12)$$

Согласно [8], при условии  $(A/\sigma_s)\cos(\phi - \hat{\phi}(h)) > 3$  возможно аппроксимировать распределение (12):

$$D(\hat{\phi}(h)) \approx \frac{A}{\sigma_s\sqrt{2\pi}} \cdot \cos u \cdot \exp\left[-\frac{A^2}{2\sigma_s^2} \sin^2 u\right]. \quad (13)$$

При условии малого отклонения фазы этот закон приближается нормальным распределением с СКО  $\sigma_\phi(h)$  и средним  $\phi(h)$ :

$$D(\hat{\phi}(h)) \approx \frac{A(h)}{\sigma_{\phi} \sqrt{2\pi}} \cdot \exp\left[-\frac{A(h)^2}{2\sigma_{\phi}^2} (\hat{\phi}(h) - \phi(h))^2\right], \quad (14)$$

$$\sigma_\phi(h) = \sigma_s / A(h) = P_0 / (\sqrt{T} \cdot A(h)). \quad (15)$$

### Статистические распределения МКФО аргументов Фурье-коэффициентов.

Существуют линейные комбинации фаз составляющих сигнала (1), не зависящие от момента начала отсчёта  $t_0$ . Такие комбинации называются межкомпонентными фазовыми отношениями (МКФО). Запишем общее выражение МКФО:

$$\begin{aligned}\Theta(k) &= \sum_{h=0}^H k(h) \cdot \hat{\phi}(h) = \sum_{h=1}^H k(h) \cdot (h\phi_0 + \theta(h) + \gamma(h)) = \\ &= \phi_0 \cdot \sum_{h=1}^H k(h)h + \sum_{h=1}^H k(h) \cdot (\theta(h) + \gamma(h))\end{aligned}\quad (16)$$

Здесь  $k \in Z$  – коэффициенты линейной комбинации,  $\gamma(h) = \hat{\phi}(h) - \phi(h)$ . Выражение (16) не зависит от  $\phi_0(h) = 2\pi \cdot hF_0 \cdot t_0$  при условии

$$\sum_{h=1}^H k(h)h = 0 \quad (17)$$

Известно, что распределение суммы  $Z$  независимых нормальных случайных величин  $X, Y$  также нормальное с дисперсией, равной сумме дисперсий исходных величин:  $D(Z) = D(X) + D(Y)$ . Кроме того,  $D(C \cdot X) = C^2 D(X)$ . Отсюда распределение  $\Theta(k)$ :

$$D(\Theta(k)) = \frac{1}{\sigma_{\Theta}(k)\sqrt{2\pi}} \cdot \exp\left[-\frac{(\Theta(k) - \Theta_0(k))^2}{2\sigma_s^2}\right], \quad (18)$$

$$\Theta_0(k) = \sum_{h=1}^H k(h) \cdot \theta(h), \quad (19)$$

при этом СКО  $\Theta(k)$  выражается через длину реализации, значения амплитуд гармоник и спектральной плотности мощности шума:

$$\sigma_{\Theta}(k) = \sqrt{\sum_{h=1}^H \sigma_{\phi}^2(h)} = \sqrt{\sum_{h=1}^H P_0^2 / (T \cdot A(h))^2} = \sqrt{P_0 / (T) \cdot \sum_{h=1}^H A^{-2}(h)} \quad (20).$$

### Результаты.

Эмпирические распределения МКФО фаз компонентов сигнала (1) были получены путём моделирования. Сигналы вида (1) генерировались при заданных значениях амплитуд, фаз, уровня шума,  $H = 3$ . Эмпирические и теоретические распределения  $\hat{\phi}(h), \Theta(k)$  при  $k = \{2, -1, 0\}$  и  $k = \{1, 1, -1\}$  приведены на рисунках 1-3 а-в соответственно. СКО теоретического распределения  $\phi(h)$  рассчитывалось по формуле (15), распределения  $\Theta(k)$  – по формуле (20).

Соответствие эмпирического и теоретического распределений проверялась при помощи теста Колмогорова-Смирнова. Суть теста заключается в оценке статистики

$$S_n = \sup_{|x| < \infty} |\hat{F}(x) - F(x)|, \quad (21)$$

где  $\hat{F}(x)$  – эмпирическая функция распределения,  $F(x)$  – теоретическая. Нулевая гипотеза не отвергается, если выполняется неравенство  $p = P(S > S_n) > \alpha$ , где  $S$  – критическое значение статистики для уровня значимости  $\alpha$ . Нулевая гипотеза была

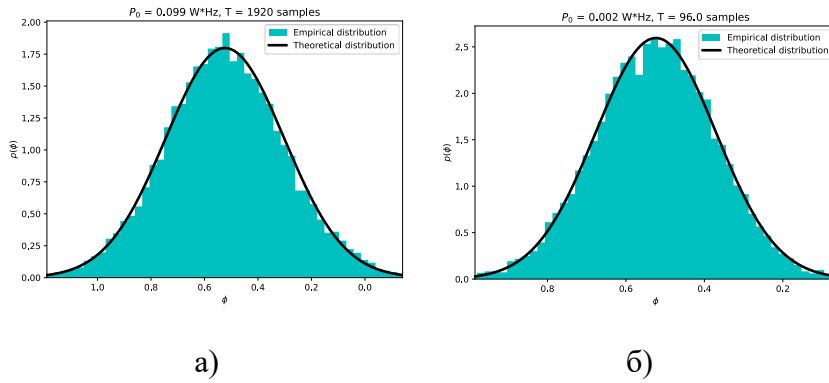


Рисунок 5 – Распределения начальной фазы первой гармоники: а)  $P_0 = 0.099$ ,  $T = 1920$ ; б)  $P_0 = 0.002$ ,  $T = 96$

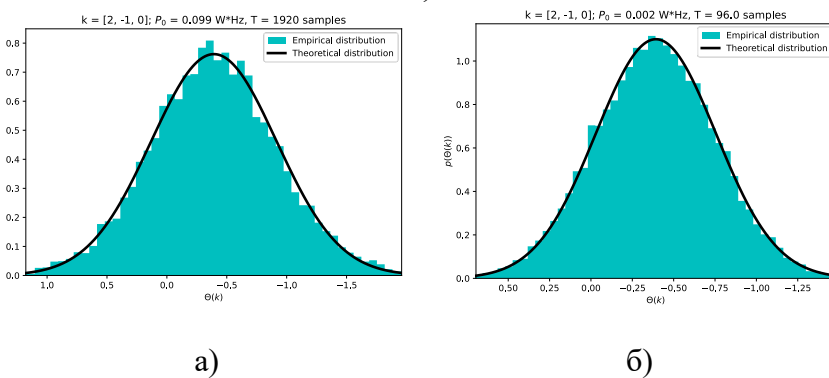


Рисунок 6 – Распределения  $\Theta(k)$ ,  $k = [2, -1, 0]$ : а)  $P_0 = 0.099$ ,  $T = 1920$ ; б)  $P_0 = 0.002$ ,  $T = 96$

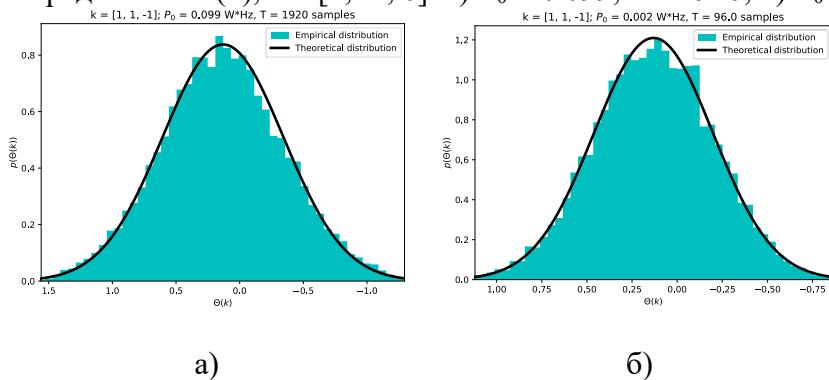


Рисунок 7 – Распределения  $\Theta(k)$ ,  $k = [1, 1, -1]$ : а)  $P_0 = 0.099$ ,  $T = 1920$ ; б)  $P_0 = 0.002$ ,  $T = 96$

сформулирована как:

- $D(\hat{\phi}(h)) = N[\phi(h), \sigma_{\phi}^2(h)]$
- $D(\Theta(k)) = N[\Theta_0(k), \sigma_{\Theta}^2(k)]$

Проверялась нулевая гипотеза при  $A = \{1, 0.8, 0.7\}$ ,  $F_0 = 10$  Гц,  $F_s = 960$  Гц. Проводилось два теста: при постоянной СПМ  $P_0 = 0.002$  Вт/Гц и  $T = [96 \dots 960]$  отсчётов,  $T = 1920$ ,  $P_0 = [0.002 \dots 0.2]$  Вт/Гц. При спектральной плотности мощности шума  $P_0 \leq 0.1$  Вт/Гц и  $T = 1920$   $p \geq 0.05$ , что не позволяет отвергнуть  $H_0$  в этих случаях. Полученный результат

свидетельствует о том, что при высоких ОСШ распределение МКФО можно считать нормальным со средним (19) и СКО (20).

#### **Заключение.**

Межкомпонентные фазовые отношения аргументов Фурье-коэффициентов сигнала вибрации, при высоких ОСШ распределены по нормальному закону. Соответствие эмпирического распределения теоретическому закону проверялась тестом Колмогорова-Смирнова. Получены аналитические выражения расчёта среднего и дисперсии МКФО, учитывающие длину реализации, значения амплитуд и уровень шума.

#### **Список литературы**

- [1] Абрамов, И.Л. Вибродиагностика энергетического оборудования / И.Л. Абрамов. – Кемерово : КузГТУ, 2011. – 81 с.
- [2] Hariharan, Dr.V. Vibration analysis of misaligned shaft – ball bearing system / Dr.V. Hariharan // Indian J. Sci. Technol. – 2009. – Т. 2. – С. 45-50.
- [3] Кечик, Д.А. Различение неисправностей оборудования при помощи межкомпонентной фазовой обработки сигналов / Д.А. Кечик, И.Г. Давыдов // Информационные радиосистемы и радиотехнологии 2022 : материалы научно-технической конференции, Минск, 29–30 ноября 2022 г. / БГУИР. – Минск : БГУИР, 2022. – С. 232-237.
- [4] Dynamic Response of a Rotating Assembly under the Coupled Effects of Misalignment and Imbalance / M. Desouki [et al.] // Shock Vib. – 2020. – Vol. 2020. – P. 26.
- [5] Kumar, A. Vibration characteristics of a rotor-bearing system caused due to coupling misalignment – a review / A. Kumar, P. Sathujoda, V. Ranjan // Vibroengineering PROCEDIA. – 2021. – Vol. 39. – P. 1-10.
- [6] Brillinger, D.R. Time Series: Data Analysis and Theory / D.R. Brillinger. – Society for Industrial and Applied Mathematics, 2001. – 564 p.
- [7] Bennett, W. Methods of Solving Noise Problems / W. Bennett // Proc. IRE. – 1956. – Vol. 44, № 5. – P. 609-638.
- [8] Левин, Б.Р. Теоретические основы статистической радиотехники / Б.Р. Левин. – Изд. 3. – М. : Радио и связь, 1989.

### **PROBABILITY DENSITIES OF INTER-COMPONENT PHASE RELATIONS**

***D.A. Kechik***

*Assistant of Department of Information Radiotechnologies, Belarusian State University of computer science and Radio Electronics, master of technical sciences  
ya.dan.kechik@yandex.by*

***I.G. Davydov***

*Associate professor of Department of Information Radiotechnologies, Belarusian State University of computer science and Radio Electronics, PhD  
davydov\_ig@bsuir.by*

*Department of Information Radiotechnologies  
Faculty of Radiotechnics and Electronics  
Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: ya.dan.kechik@yandex.by*

**Abstract.** Approximation of inter-component phase relations (ICPR) of components of corrupted by additive gaussian noise polyharmonic signal has been presented in this paper. ICPR probability distribution law has been shown to be close to normal.

**Key words:** inter-component phase relations, digital signal processing, statistics, Kolmogorov-Smirnov test.

УДК 004.021:004.75

## РАБОТА АЛГОРИТМА КОЛЛАПСА ВОЛНОВОЙ ФУНКЦИИ НА ПРИМЕРЕ ИГРЫ «TOWER OF ELEVATION»



**С.Ю. Слюсарь**

Учащийся Учреждения образования «Национальный детский технопарк», учащийся ГУО «Гимназия г. Логойска»



**Д.С. Коваленко**

Учащийся Учреждения образования «Национальный детский технопарк», учащийся ГУО «Гимназия г. Логойска»



**М.С. Ильясова**

магистрант кафедры инженерной психологии и эргономики БГУИР



**Ф.В. Усенко**

магистрант кафедры инженерной психологии и эргономики БГУИР



**Л.Р. Коркин**

магистр технических наук, ассистент кафедры инженерной психологии и эргономики БГУИР



**А.М. Прудник**

доцент кафедры инженерной психологии и эргономики БГУИР, кандидат технических наук, доцент  
[aleksander.prudnik@bsuir.by](mailto:aleksander.prudnik@bsuir.by)

### **С.Ю. Слюсарь**

Обучается в гимназии г. Логойска. Область научных интересов связана с технологиями виртуальной и дополненной реальности.

### **Д.С. Коваленко**

Обучается в гимназии г. Логойска. Область научных интересов связана с технологиями виртуальной и дополненной реальности.

### **М.С. Ильясова**

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с автоматизированным тестированием информационных систем.

### **Ф.В. Усенко**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой манипуляторов для дистанционного управления.

### **Л.Р. Коркин**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с системами распознавания снимков.

### **А.М. Прудник**

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с взаимодействием человека с компьютером, интерфейсами информационных систем, пользовательскими интерфейсами, front-end web development.

**Аннотация.** В статье описывается алгоритм коллапса волновой функции и пример его работы на примере игры «Tower of elevation».

**Ключевые слова:** Алгоритм коллапса волновой функции, игра «Tower of elevation», Unity, MagicaVoxel, тайл, мир, карта, Visual Studio.

### **Введение.**

Алгоритмы имеют большую популярность в «мире» разработчиков. Их используют в играх разного масштаба: от совсем коротеньких и с относительно маленькой картой, до бесконечных миров, где игрок вряд ли сможет дойти до границы мира.

Цели доклада:

1. Рассказать о алгоритме коллапса волновой функции;
2. Показать работу алгоритма на примере разработанной игры;
3. Рассказать о плюсах и минусах алгоритма.

### **Основная часть.**

В настоящее время разработка игр является одним из наиболее распространенных направлений в компьютерной индустрии. Некоторые из них создаются просто путем ручной расстановки элементов карты дизайнерами, другие ограничиваются небольшой локацией, в которой игрок находится, а третьи используют компьютер для создания огромных миров. Мы относимся к последней категории разработчиков, что особенно полезно для игр с большой картой, генерируемой в произвольном порядке, имеющей множество вариаций. Примером такой игры является «Minecraft» [1], где большое количество вариаций блоков расставляется в произвольном порядке в соответствии с используемым алгоритмом. Практически все эти алгоритмы работают по принципу, который заключается в сопоставлении границы модели с другой моделью, чтобы получить единый и гармоничный результат.

### **Описание алгоритма коллапса волновой функции.**

Алгоритм коллапса волновой функции (Wavefunction Collapse Algorithm) учит компьютер «импровизировать». На входе задаются архетипичные данные и создаются процедурно-генерируемые данные, похожие на исходные и приведенные на рисунке 1. Чаще всего алгоритм используется для создания изображений, но может также строить города, скейтпарки и другие объекты. В нашей игре алгоритм строит локации, на которых происходят все действия игрока. Каждая часть локации была создана не человеком, а алгоритмом. Это означает, что игроку будет интересно исследовать мир игры, так как он не сможет предугадать, где что находится [2].

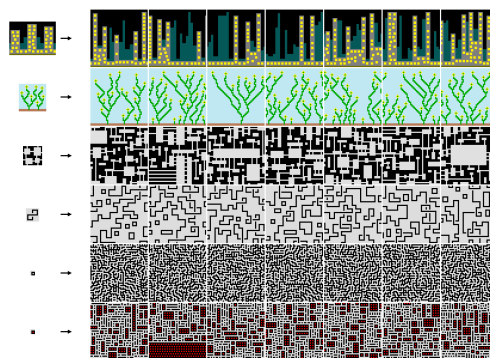


Рисунок 1. Создание процедурно-генерируемых данных, похожих на исходные

Коллапс волновой функции — это очень «независимо мыслящий» алгоритм, не требующий практически никакой помощи или инструкций извне. Необходим только пример стиля, которого нужно достичь, а всё остальное алгоритм сделает сам. Несмотря на свою самодостаточность, алгоритм является простым и не использует никаких нейронных сетей, случайных лесов или чего-то другого, похожего на машинное обучение [3].

### Реализация работы алгоритма на примере игры «Tower of elevation».

Игра «Tower of Elevation» была создана для исследования работы алгоритма коллапса волновой функции. В процессе создания игры были выполнены следующие этапы:

1. Создание в MagicaVoxel тайлов, соответствующих требованиям алгоритма коллапса волновой функции.

2. Разработка алгоритма в среде программирования Visual Studio.

3. Интеграция алгоритма с ранее разработанными тайлами и тестирование его работы.

4. Анализ плюсов и минусов работы алгоритма коллапса волновой функции.

Следует рассмотреть каждый этап более подробно.

Первым этапом являлось создание тайлов, которые должны были соответствовать требованиям алгоритма коллапса волновой функции. Для этого каждый тайл должен был иметь определенную границу, чтобы алгоритм мог взаимодействовать с ними [4]. Если границы у каждого тайла разные, то алгоритм не сможет их обрабатывать. На рисунке 2 показаны все тайлы, которые были использованы в работе.

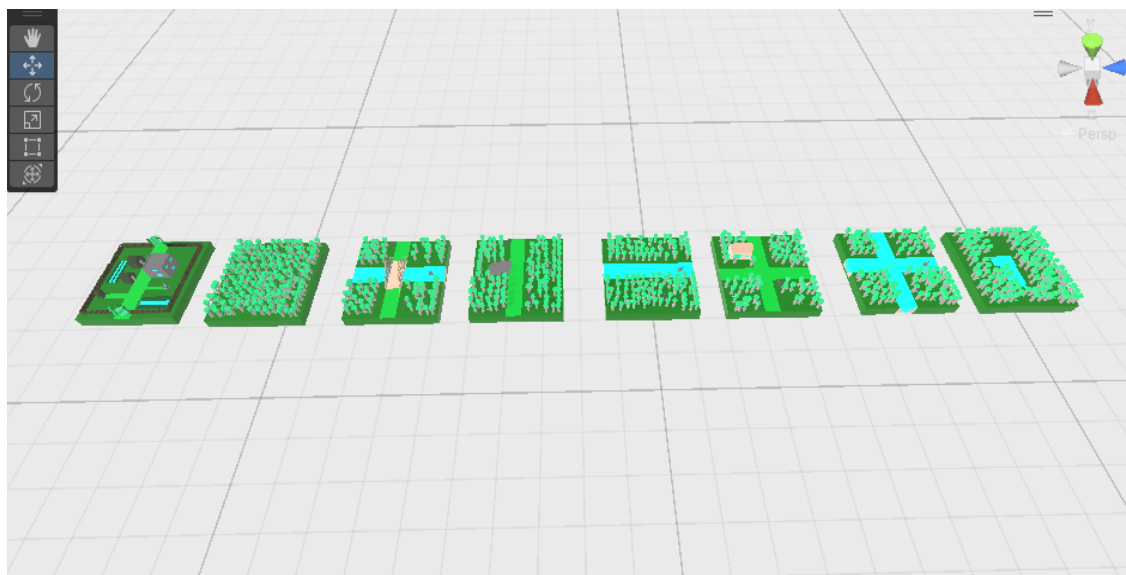


Рисунок 2. Виды тайлов, которые были использованы в работе разработки игры

Затем приступили к разработке скриптов в интегрированной среде разработки Visual Studio. Написанный программный код предназначался для тайлов и места их расположения. Примеры начала написания кода продемонстрированы на рисунках 3 и 4, а фрагмент кода приведен ниже.

Скрипт, представленный на рисунке 3, выполняет роль «конструктора». То есть скрипт считывает из каких цветов состоит боковая сторона определённого тайла и потом благодаря этому скрипту программа может соединять различные тайлы логически и без резких изменений для глаза человека. Одним словом, с помощью этого скрипта делаются плавные и логические переходы между разными тайлами.



```
public void CalculateSideColors()
{
    ColorsRight = new byte[TilesSideVoxels * TilesSideVoxels];
    ColorsForward = new byte[TilesSideVoxels * TilesSideVoxels];
    ColorsLeft = new byte[TilesSideVoxels * TilesSideVoxels];
    ColorsBack = new byte[TilesSideVoxels * TilesSideVoxels];
    for (int y = 0; y < TilesSideVoxels; y++)
    {
        for (int i = 0; i < TilesSideVoxels; i++)
        {
            ColorsRight[y * TilesSideVoxels + i] =
GetVoxelColor(verticalLayer y, horizontalOffset i, Direction.right);
            ColorsForward[y * TilesSideVoxels + i] =
GetVoxelColor(verticalLayer y, horizontalOffset i, Direction.forward);
            ColorsLeft[y * TilesSideVoxels + i] =
GetVoxelColor(verticalLayer y, horizontalOffset i, Direction.left);
            ColorsBack[y * TilesSideVoxels + i] =
GetVoxelColor(verticalLayer y, horizontalOffset i, Direction.back);
        }
    }
}
```

Рисунок 3. Скрипт для плавного и логического перехода между тайлами

Скрипт, представленный на рисунке 4, устанавливает определённые координаты для расположения тайлов. То есть при задании определённых координат алгоритм сам решает, какой тайл из списка тайлов подходит в определенное положение на сцене.

```
public void PlaceTile(int x, int y)
{
    List<VoxelTile> availableTiles = new List<VoxelTile>();
    foreach (VoxelTile tilePrefab in TilePrefabs)
    {
        if (CanAppendTile(existingTile: spawnedTiles[x - 1, y],
tileToAppend: tilePrefab, Direction.Left) &&
            CanAppendTile(existingTile: spawnedTiles[x + 1, y],
tileToAppend: tilePrefab, Direction.Right) &&
            CanAppendTile(existingTile: spawnedTiles[x, y - 1],
tileToAppend: tilePrefab, Direction.Back) &&
            CanAppendTile(existingTile: spawnedTiles[x, y + 1],
tileToAppend: tilePrefab, Direction.Forward))
        {
            availableTiles.Add(tilePrefab);
        }
    }
    if (availableTiles.Count == 0) return;
    VoxelTile selectedTile = GetRandomTile(availableTiles);
    Vector3 position = new Vector3(x, y: 0, z: y) * selectedTile.VoxelSide
* selectedTile.TileSideVoxels;
    spawnedTiles[x, y] = Instantiate(selectedTile, position,
Quaternion.Identity);
}
}
```

Рисунок 4. Скрипт для автоматического расположения тайлов по координатам

И заключительным этапом разработки игры является применение алгоритма. Результат выполнения алгоритма показан на рисунке 5.

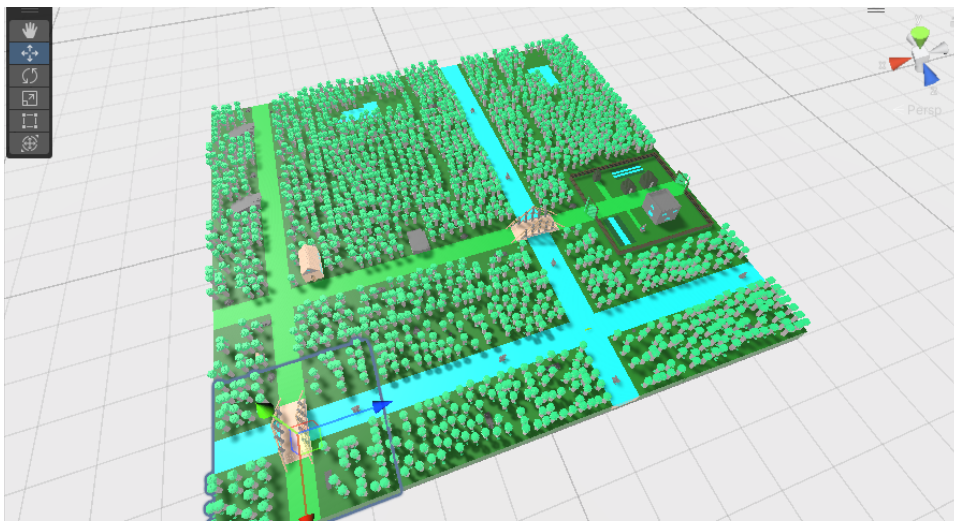


Рисунок 5. Результат выполнения алгоритма коллапса волновой функции

Перечислим следующие преимущества алгоритма коллапса волновой функции:

1. Простота работы алгоритма, что делает его привлекательным для использования.
2. Удобство в использовании, что упрощает процесс его применения..
3. Быстрота достижения результата действий алгоритма и возможность его перезапуска с дополнительными настройками.

К недостаткам алгоритма коллапса волновой функции можно отнести:

1. Ограниченность возможностей создаваемых объектов.
2. Необходимость точного определения границ тайлов.
3. Ограниченность разнообразия создаваемых объектов.
4. Необходимость многочисленных экспериментов для определения наиболее подходящих правил и параметров алгоритма.
5. Сложность создания и отладки алгоритма для новых типов объектов.

### **Заключение.**

В результате выполнения проекта разработана игра виртуальной реальности «Tower of elevation» с генерацией карты при помощи алгоритма коллапса волновой функции. Игра рассчитана на аудиторию до 18 лет, хотя в неё может играть и взрослый человек. Так как приложение представляет собой игру с некоторым сюжетом, пользователю будет интересно «продвигаться» по игре, разыскивая пасхалки, сделанные разработчиками. Так как в игре после начального тестирования не обнаружено никаких багов и всё работает правильно, то можно начать крупное тестирование, набирая бета-тестеров и разрабатывать обновления для игры.

### **Список литературы**

- [1] Minecraft [Электронный ресурс]. URL: <https://www.minecraft.net/ru-ru>. (Дата обращения: 12.04.2023).
- [2] Генерация уровня из 3D тайлов: часть 1 [Электронный ресурс]. URL: [https://www.youtube.com/watch?v=iG\\_n23W952Y&list=RDCMUC6wnai488mwec\\_FUVfdl84w&index=10](https://www.youtube.com/watch?v=iG_n23W952Y&list=RDCMUC6wnai488mwec_FUVfdl84w&index=10). (Дата обращения: 12.04.2023).

[3] Генерация уровня из 3D тайлов: часть 2 [Электронный ресурс]. URL: <https://www.youtube.com/watch?v=fMtbvWCGke4&t=636s>. (Дата обращения: 12.04.2023).

[4] Генерация уровня из 3D тайлов: часть 3 [Электронный ресурс]. URL: [https://www.youtube.com/watch?v=EG0iERtXc0U&list=RDCMUC6wnai488mwec\\_FUVfdl84w&index=1](https://www.youtube.com/watch?v=EG0iERtXc0U&list=RDCMUC6wnai488mwec_FUVfdl84w&index=1). (Дата обращения: 12.04.2023).

## **OPERATION OF THE WAVE FUNCTION COLLAPSE ALGORITHM ON THE EXAMPLE OF THE GAME "TOWER OF ELEVATION"**

**S.Y. Slusar**

*Student of the educational Institution "National Children's Technopark", student GUO "Gymnasium of Logoysk"*

**D.S. Kovalenko**

*Student of the educational Institution "National Children's Technopark", student GUO "Gymnasium of Logoysk"*

**M.S. Ilyasova**

*Master's student of the Department of Engineering Psychology and Ergonomics of BSUIR*

**F.V. Usenko**

*Master's student of the Department of Engineering Psychology and Ergonomics of BSUIR*

**L.R. Korkin**

*Master of Technical Sciences, Assistant of the Department of Engineering Psychology and Ergonomics of BSUIR*

**A.M. Prudnik**

*Associate Professor of Engineering Psychology and Ergonomics Department of BSUIR, Candidate of Technical Sciences, Associate Professor [aleksander.prudnik@bsuir.by](mailto:aleksander.prudnik@bsuir.by)*

*Department of Engineering Psychology and Ergonomics*

*Faculty of Computer-Aided Design*

*Belarusian State University of Informatics and Radio Electronics, Republic of Belarus*

*E-mail: [aleksander.prudnik@bsuir.by](mailto:aleksander.prudnik@bsuir.by)*

УДК 378.141 : 004.9

## ПОВЫШЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ПОДАЧИ ЗАЯВЛЕНИЙ И ЗАЧИСЛЕНИЯ ЗА СЧЕТ МАСШТАБИРОВАНИЯ БАЗЫ ДАННЫХ



**Б. В. Никульшин**  
Заведующий кафедрой  
электронных вычислительных  
машин,  
Белорусского  
государственного  
университета информатики и  
радиоэлектроники кандидат  
технических наук,  
доцент  
nik@bsuir.by



**В.М. Бондарик**  
Декан факультета  
доуниверситетской  
подготовки и  
профессиональной  
ориентации  
Белорусского  
государственного  
университета информатики и  
радиоэлектроники,  
кандидат  
технических наук,  
доцент  
bondarik@bsuir.by



**Н. В. Русина**  
Ассистент кафедры  
программного обеспечения  
информационных технологий,  
кафедры физики Белорусского  
государственного  
университета информатики и  
радиоэлектроники  
rusina@bsuir.by

### **Б. В. Никульшин**

Образование: Минский радиотехнический институт. Область научного интереса – инновационные технологии принятия решений в проектной и управленческой деятельности, методы и модели принятия решений, теория системного анализа и принятия решений в инфокоммуникациях, системный анализ.

### **В.М. Бондарик**

Образование: 1983-1988 – Минский радиотехнический институт, специальность «Конструирование и производство радиоаппаратуры», квалификация – инженер-конструктор-технолог. Область научного интереса – проектирование медицинских электронных систем, внедрение дистанционных образовательных технологий.

### **Н. В. Русина**

Образование : МГУ им. А. Кулешова, специальность-математик, физик. Область научных интересов связана с разработкой информационной системы электронного зачисления в ВУЗы, организацией учебного и научно-исследовательского процессов в техническом университете.

**Аннотация.** Описан один из возможных вариантов масштабирования базы данных с использованием шардинга для увеличения скорости обработки структурированных данных и производительности всей системы. Определены большинство возможностей и ограничений шардинга. Описано применение его на практике в приложении.

**Ключевые слова:** масштабирование, шардинг, база данных, производительность.

Организация конкурса, зачисление в учреждение высшего образования и передача данных личного дела студента в информационную систему университета – финальный этап организации приемной кампании учреждения высшего образования.

Приемная кампания для получения общего высшего и специального высшего образования в Республике Беларусь регламентируется следующими основными нормативно-правовыми актами:

1. Правила приема лиц для получения общего высшего и специального высшего образования, утвержденные Указом Президента Республики Беларусь от 27.01.2022 № 23.
2. Порядок приема в учреждение высшего образования на текущий год.
3. Положение о приемной комиссии учреждения высшего образования, утвержденное Постановлением Министерства образования Республики Беларусь от 23.03.2006 № 23.

Организация приемной кампании для крупных учреждений высшего образования без применения средств автоматизации неэффективна.

Белорусский государственный университет информатики и радиоэлектроники (БГУИР) уже более десяти лет ведет приемную кампанию с применением автоматизированной системы подачи заявлений и зачисления (АСПЗиЗ) [1]. При этом автоматизированная система ежегодно дорабатывается, так как регулярно вносятся изменения в порядок приема учреждения образования и в правила приема лиц для получения общего высшего и специального высшего образования, а также углубленного высшего образования.

Приемная кампания в 2023 году проходит в несколько этапов:

- прием документов и зачисление в университет льготных категорий абитуриентов и абитуриентов, планирующих получение высшего образования на условиях целевой подготовки – первая половина июня;
- прием документов и зачисление в университет лиц для получения углубленного высшего образования – конец июня – первая половина июля;
- прием документов и зачисление в университет лиц для получения общего высшего и специального высшего образования за счет средств республиканского бюджета – вторая половина июля;
- прием документов и зачисление в университет лиц для получения общего высшего и специального высшего образования на условиях оплаты обучения из собственных средств – первая половина августа.

При автоматизации приемной кампании необходимо неоднократно фильтровать и пересылать большое количество данных различных типов. Существует несколько способов повышения производительности запросов к базам данных [2].

Можно повышать производительность АСПЗиЗ за счет вертикального масштабирования. Это экстенсивный путь: необходимо обновить оборудование для сервера, добавить количество оперативной памяти и т.д. Он является достаточно дорогим, длительным по времени имеет свой предел роста. Можно приобрести лучшее оборудование, однако и оно может не справиться со всеми требованиями приложения.

При горизонтальном масштабировании предполагается расширение вычислительных ресурсов, доступных приложению, за счет увеличения количества серверов, на которых размещена база данных. При размещении базы данных на одном сервере в какой-то момент она перестает справляться с нагрузкой и значительно увеличивается время отклика приложения.

Для ускорения работы АСПЗиЗ в условиях массовых запросов пользователей к базам данных приложения можно использовать стратегию шардинга, которая позволяет распределять данные между разными физическими серверами. Это существенно облегчает обработку данных [3].

Существует два типа шардинга: вертикальный и горизонтальный. При организации приемной кампании с применением АСПЗиЗ предлагается использовать горизонтальный шардинг – разделение базы данных на разные сервера. Этот способ рекомендуется использовать при обработке больших таблиц, которые не умещаются на одном сервере и замедляют работу приложения. Разделение базы данных делается по следующему алгоритму:

- на нескольких новых серверах создается одна и та же база данных (только структура, без данных);

- определяются таблицы, данные которых должны присутствовать во всех базах данных, и те, у которых данные будут уникальными во всей системе;

- перед каждым обращением к базе данных происходит выбор нужного соединения.

При реализации этого алгоритма необходимо произвести классификацию таблиц, что позволяет в дальнейшем в программном коде добавить процесс синхронизации и переноса данных. При этом заранее известно с каким шардом работает приложение и нет необходимости менять соединение между главным шардом по умолчанию и дополнительным шардом.

Процесс переноса данных между шардами в реальном времени может быть длительным из-за большого количества данных, что уменьшит скорость работы как сервера, с которого будут переноситься данные, так и сервера, на который они будут мигрировать.

Перед процессом переноса данных необходимо заблокировать все операции на таблицах, данные которых будут переноситься на другой шард. Для сокращения времени передачи данных рекомендуется создать специальный процесс, генерирующий в файловой системе SQL файлы, в которых будут находиться все данные для импорта, а также создать новый процесс, который займется импортом этих данных. Все эти действия рекомендуется выполнять на специальном сервере, чтобы не замедлять работу самого приложения.

Недостатком горизонтального шардинга является необходимость организации поиска и фильтрации данных. Данная проблема может быть решена путем написания дополнительного кода, в котором производится выборка данных со всех шардов, затем данные группируются в один массив, и в итоге в полученном массиве данных производится внутренняя фильтрация данных.

Вторым недостатком горизонтального шардинга может стать уникальность данных, однако эта проблема решается довольно просто. На этапе разработки архитектуры приложения необходимо заранее спланировать возможную нагрузку и на основе полученных результатов можно определить количество серверов для комфортной работы приложения. Для устранения ситуации, когда две или более записи с разных шардов имеют одинаковый ID, рекомендуется произвести следующие операции с серверами баз данных:

- на каждом сервере баз данных нужно установить начальное значение ID, т.е. на первом сервере это будет 1, на втором – 2, и т.д.;

- определить значение `auto_increment` величине, равной количеству планируемых серверов.

Преимуществом горизонтального масштабирования является возможность практически бесконечного масштабирования разрабатываемой базы данных.

Процесс горизонтального масштабирования базы данных является эффективным способом как повышения общей производительности системы и обработки больших объемов данных, так и снижения общей нагрузки на один единственный сервер. Предложенный способ может быть применен как при планировании архитектуры приложения, так и на готовом проекте. Однако применение горизонтального масштабирования на готовом проекте рекомендуется применять только в крайнем случае при снижении производительности приложения до критического уровня.

При разработке АСПЗиЗ было применено горизонтальное масштабирование с использованием шардинга, т.к. на стадии планирования приложения было решено изначально предотвратить возможное уменьшение производительности приложения. Разработанное программное обеспечение успешно работает в сетевой версии со множеством клиентов.

Таким образом, в БГУИР постоянно совершенствуется единое информационное пространство поддержки приемной кампании, что позволит в комфортных условиях принимать в БГУИР наиболее подготовленных абитуриентов при одновременном сокращении временных и финансовых затрат.

#### **Список литературы**

[1] Развитие единого информационного пространства приемной кампании Белорусского государственного университета информатики и радиоэлектроники / Никульшин Б. В. [и др.] // Дистанционное обучение – образовательная среда XXI века : материалы XI Международной научно-методической конференции, Минск, 12-13 декабря 2019 г. / редкол.: В. А. Прытков [и др.]. – Минск : БГУИР, 2019. – С. 219.

[2] Повышение быстродействия системы электронного обучения / Бондарик В. М. [и др.] // Глобальный научный потенциал. – 2021. – № 9(126). – С. 66–69.

[3] Слепов, И. В. Масштабирование базы данных с использованием шардинга / И. В. Слепов, А. Д. Тюменцев, В. М. Бондарик // BIG DATA Advanced Analytics: collection of materials of the fourth international scientific and practical conference, Minsk, Belarus, May 3 – 4, 2018 / editorial board: M. Batura [etc.]. – Minsk, BSUIR, 2018. – P. 184–186.

## **IMPROVING PERFORMANCE AUTOMATED APPLICATION AND ENROLLMENT SYSTEM BY SCALING A DATABASE**

***Б. В. Никульшин***

*Head of the Department of Electronic Computing Machines, Belarusian State University of Informatics and Radioelectronics, candidate of technical Sciences*

***В.М. Бондарик***

*Dean of the faculty of Pre-University training and vocational guidance Belarusian State University of Informatics and Radioelectronics, candidate of technical*

***Н. В. Русина***

*Assistant of IT Software Department, Department of Physics, Belarusian State University of Informatics and Radioelectronics*

*Department of Electronic Computing Machines*

*Department of Physics*

*Faculty of Pre-University training and vocational guidanc*

*Belarusian State University of Informatics and Radioelectronic*

*E-mail: nik@bsuir.by*

**Abstract.** The main goal of this research is to describe one of the possible ways of scaling a database using sharding to increase the processing speed of structured data and, in general, the performance of the entire system. During the research, I was able to determine most of the possibilities and limitations of sharding, and also applied it in practice in a application.

**Keywords:** scaling, sharding, database, performance.

УДК 004.021

## BIG DATA КАК ОСНОВА ПОСТРОЕНИЯ ЭФФЕКТИВНОЙ СИСТЕМЫ МОНИТОРИНГА ПАНДЕМИЙ НА ПРИМЕРЕ COVID-19



**Д.В. Орлов**  
Студент 4 курса,  
кафедра ЭВМ, БГУИР  
dim4work@gmail.com



**С.Н. Нестеренков**  
Кандидат технических  
наук, доцент,  
декан факультета  
компьютерных систем и  
сетей  
s.nesterenkov@bsuir.by



**Д. А. Жалейко**  
инженер-программист  
кафедра ЭВМ, БГУИР  
a.n.markov@bsuir.by

### **Д.В. Орлов**

Студент 4 курса специальности “Вычислительные машины, системы и сети” БГУИР.

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

### **Д.А. Жалейко**

Окончил БГУИР в 2021 году по специальности “Вычислительные машины системы и сети”

**Аннотация.** После вспышки коронавируса (SARS-CoV-2) в 2019 году он заразил миллионы людей и унес жизни десятки тысяч человек. Во время пандемии коронавируса, применение Big Data стало одним из немногих мощных средств для борьбы с вирусом. В настоящее время многие страны и исследовательские учреждения используют Big Data для отслеживания и контроля распространения инфекционного заболевания. В будущем люди смогут использовать возможности Big Data для борьбы с такими эпидемиями. К примеру сравнительное геномное исследование вариантов вируса, ускоренное с помощью Big Data, может дать важную информацию о мутациях вируса. Эта статья поведает о применении Big Data во время пандемии COVID-19 и как бороться с подобными эпидемиями в будущем. Много приложений и программных решений было создано с использованием этих данных, отслеживающих и предсказывающих поведение инфекции. Следовательно такие приложения и методы обработки помогут человечеству в борьбе с похожими случаями в будущем.

**Ключевые слова:** Big Data, COVID-19, эпидемия, общество.

### **Введение.**

Пандемия COVID-19 стала беспрецедентной ситуацией в области общественного здравоохранения. Эта болезнь стала серьезным вызовом эпидемиологическому планированию и системе здравоохранения.

Этот инцидент в области общественного здравоохранения, охвативший весь мир, не является первым в истории человечества; подобные случаи происходили несколько раз за последние сотни лет. Черная смерть, которая бушевала с 1346 по 1353 года, унесла жизни около 200 миллионов людей. Пандемия холеры с 1852 по 1860 годы унесла около 1 миллиона жизней. Тем не менее с развитием медицина и института здравоохранения появление такой эпидемии как Коронавирус, предупреждает мир, что он не готов к таким случаям.



По данным Всемирной Организации Здравоохранения, по состоянию на 30 августа 2022 года, 216.303.376 случаев COVID-19 было подтверждено во всем мире, и число погибших достигло 4.498.451. Среди них США имеет наибольшее количество подтвержденных случаев, Индия идет следующей. Хотя более 4 миллиардов вакцин было выпущено, это никак не изменило ситуацию из-за нового варианта вируса Delta.

В настоящее время глобальные цели по борьбе с вирусом сосредоточены на сведении к минимуму его распространения и смертности, а также минимизировать ущерб обществу. Поэтому многие научно-исследовательские институты и государственные учреждения работают сообща и используют передовые технологии: анализ Big Data и искусственный интеллект для борьбы с такими пандемиями. Например США и Шанхай используют преимущества анализа Big Data для борьбы с коронавирусом. Ключевая информация, такая как физическое состояние и история путешествий более 100.000 человек была записана, чтобы справиться с пандемией. Эта статья направлена на применяемые технологии Big Data в условиях коронавируса. Узнайте, как использовать эти инструменты, чтобы позволить человечеству лучше реагировать на такие вспышки в будущем [1].

### Приложения использующие Big Data для исследования COVID-19

Некоторые исследователи и правительственные учреждения используют Big Data для отслеживания случаев заражения в реальном времени. Анализ эпидемии включает в себя все данные такие как положительные случаи заражения, смерти, людей выздоровевших от болезни, истории поездок, плотность населения, контакты с инфицированными. С применением искусственного интеллекта и машинного обучения, данные могут быть обработаны для построения модели болезни, которая может предсказать уровень заражения (высокий или низкий) и его влияние [2]. AarogyaSetu, мобильное приложение для слежения, запущенное правительством Индии, является отличным примером того, как Big Data можно использовать для борьбы с болезнью. Рабочий процесс этого приложения представлен на рисунке.



Рисунок 1. Рабочий процесс AarogyaSetu

Собирая данные о местоположении через GPS или Wi-Fi, а затем сравнивая его с существующей базой данных, приложение может выявить людей, которые могли контактировать с людьми, у которых подтвержденный случай заболевания, что обеспечивает раннее обнаружение и отслеживание цепочки заражения.

Кроме того Big Data также применяется для диагностики и прогноза COVID-19. Поскольку радиологическое проявление инфекции ковид и вирусной пневмонии очень

похожи, различие их в клинических условиях становится очень важным заданием. Таким образом КТ грудной клетки пациентов инфицированных COVID-19 и больных вирусной пневмонией используются в качестве входных данных системы глубокого обучения. В итоге финальная точность различия этих болезней составила 86%, а уровень точности в различии пациентов с ковид и здоровых составила 94%. Машинное обучение также приняло участие при разработке алгоритма прогнозирования.

Алгоритм используется для прогнозирования риска смертности инфицированных людей. Это помогает выявить пациентов с потенциальной потребностью в дальнейшей медицинской помощи, что позволяет медицинскому персоналу рационально распределять ресурсы больницы.

### Будущее применение Big Data при пандемиях

Несмотря на некоторый прогресс в разработке вакцин и терапевтических средств, COVID-19 остается серьезной угрозой для здоровья населения, и продолжается исследование и внедрение новых методов борьбы с этой инфекцией. Исследователи разработали инструменты и методы из прошлого опыта, который может быть полезен в борьбе как с COVID-19 так и другими подобными заболеваниями в будущем. В этом разделе будут описаны четыре основные области применения, включая отслеживание, предсказание, диагностику и лечение.

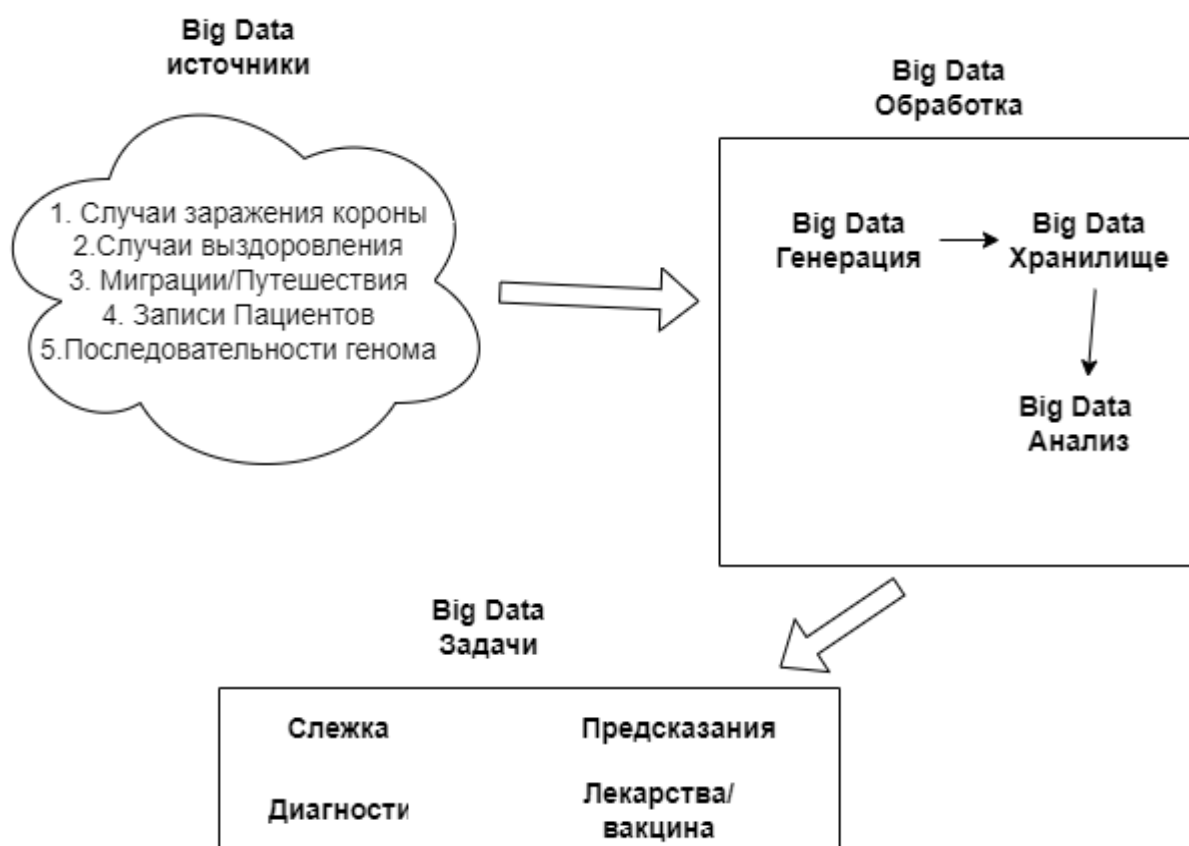


Рисунок 2. Борьба с инфекциями используя Big Data

Одна из особенностей COVID-19 является то, что больные могут передавать вирус до того, как у них появятся какие-либо симптомы. Следовательно вирус становится очень заразным, это критично для правительства, чтобы отслеживать и сдерживать инфекцию, а также принимать своевременные меры для борьбы.

Помимо отслеживания и прогнозирования распространения вируса, Big Data также могут применяться для диагностики и лечения. Надежная, быстрая и удобная диагностика необходима для борьбы с вирусом.

В условиях пандемии это может спасти большое количество жизней, но и ограничить его распространение; большое количество медицинских данных могут выступать в качестве обучающего набора для машинного обучения. Прежде всего, применение Big Data может значительно сократить время обнаружения болезни. Так как для определения точного исследования нужно обладать большим количеством оборудования, денег и времени. Следовательно, у многих стран нету возможностей, чтобы эффективно противостоять вирусу.

Мобильное приложение основанное на Big Data: AI4COVID-19, которое можно использовать для постановки диагноза на основании кашля.

Точность этого приложения при диагностике достигает 90%, что делает его удобным, дешевым и эффективным средством для борьбы с вирусом, которое может позволить любое медицинское учреждение.

Big Data широко используется в исследованиях лекарственных средств. Вакцины, таблетки и т.п., разработанные на основе этих данных, могут сыграть решающую роль в реагировании на подобные эпидемии. Google DeepMind - компания известная известная благодаря программе для игры в го AlphaGo, теперь может использовать искусственный интеллект для предсказания мембраны SARS-CoV-2 белков, которые могут быть полезны для разработки новых лекарств [3-4].

### **Заключение**

Были рассмотрены и представлены механизмы контроля и профилактические меры, которые люди использовали до сих пор, включающие в себя Big Data и инструменты работы над ними, в условиях глобального кризиса вызванного COVID-19 пандемией. Также исследуются будущее развитие этого направления и инструменты для чрезвычайного реагирования на подобные эпидемии.

Во-первых, Big Data можно использовать для разработки приложений отслеживания и прогнозирования болезни. Затем более быстрый и безопасный клинический диагноз может быть поставлен с помощью искусственного интеллекта на основе большого кластера данных.

Во-вторых, данные Big Data могут быть задействованы при изучении вируса, и поисков лекарства от него. Хотя уже сейчас видны положительные результаты применения этих технологий для борьбы с коронавирусом, методы обработки данных находятся на относительно ранней стадии. Следовательно, в данном случае трудно судить о том, какой алгоритм больше подходит для обнаружения вирусов. Другой вопрос - конфиденциальность и безопасность персональных данных людей. Во время пандемии COVID-19 правительства просили граждан делиться личной информацией такой как: местоположение, история путешествий, отчеты тестов и т.д.

Однако люди часто не желают делиться своими личными данными, потому что они заботятся о ее конфиденциальности.

Более того Big Data могут содержать дезинформацию, что несомненно имело негативное влияние на борьбу с пандемией.

Несмотря на то что имеется ряд проблем, это не отменяет того факта что Big Data во время пандемии является мощным оружием для людей в борьбе с болезнью. В будущем нужно проводить исследования в этой области для решения проблем на этапе сбора данных и их обработки.

### Список литературы

- [1] Big Data: Concepts, Technology and Architecture / В. Balusamy [и др.]. – Hoboken : John Wiley & Sons, Inc., 2021. – 368 с.
- [2] Big-Data Analytics and Cloud Computing: Theory, Algorithms and Applications / М. Trovati [и др.]. – Berlin : Springer International Publishing, 2015. – 185 с.
- [3] Нестеренков, С.Н. Применение больших данных в электронном образовании / С.Н. Нестеренков, М.И. Макаров, Н.В. Ющенко, А.Д. Радкевич // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов V Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 13-14 марта 2019 года). В 2 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : БГУИР, 2019. - С. 242-245.
- [4] Беляк, А. А. Анализ производительности технологии Hadoop / А. А. Беляк, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 19-20 мая 2021 года): / редкол. : В. А. Богуш [и др.]. – Минск : Бестпринт, 2021. – С. 343–346.
- [5] Security and Privacy for Big Data, Cloud Computing and Applications / W. Ren [и др.]. – London : The Institution of Engineering and Technology, 2019. – 328 с.
- [6] Калоша, А.Л. Система анализа качества текстовых коллекций / А.Л. Калоша, М.А. Медунецкий, М.П. Хоронко, А.А. Александров, А.И. Гридасов, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VI Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 20-21 мая 2020 года): в 3 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : Бестпринт, 2020. - С. 369-375.

## BIG DATA AS A BASIS FOR BUILDING AN EFFECTIVE PANDEMIC MONITORING SYSTEM ON THE EXAMPLE OF COVID-19

**D.V. ORLOV**

*Pregraduated student of the  
BSUIR,  
department of electronic  
computers*

**S.N. NESTERENKOV**

*PhD, Associate professor, Dean  
of the Faculty of Computer  
Systems and Networks*

**D.A. ZHALEIKO**

*Software engineer  
BSUIR,  
department of electronic  
computers*

*Center for Informatization and Development of the Belarusian University of State Informatics and Radioelectronics, Republic of Belarus.*

*E-mail: [dimm4work@gmail.com](mailto:dimm4work@gmail.com), [s.nesterenkov@bsuir.by](mailto:s.nesterenkov@bsuir.by), [a.n.markov@bsuir.by](mailto:a.n.markov@bsuir.by)*

**Abstract.** Since the outbreak of the coronavirus (SARS-CoV-2) in 2019, it has infected millions and claimed the lives of tens of thousands. During the coronavirus pandemic, the use of Big Data has become one of the few powerful tools to fight the virus. Currently, many countries and research institutions use Big Data to track and control the spread of an infectious disease. In the future, people will be able to use the power of Big Data to fight such epidemics. For example, a comparative genomic study of virus variants, accelerated with the help of Big Data, can provide important information about virus mutations. This article will tell you about the use of Big Data during the COVID-19 pandemic and how to deal with similar epidemics in the future. Many applications and software solutions have been created using this data to track and predict infection behavior. Therefore, such applications and processing methods will help humanity in the fight against similar cases in the future.

**Keywords.** Big Data, COVID-19, epidemic, society

УДК 004.02:81-13

## РАЗРАБОТКА ИНСТРУМЕНТАЛЬНЫХ СРЕДСТВ ДЛЯ ИЗУЧЕНИЯ И АНАЛИЗА УРОВНЯ ВЛАДЕНИЯ ИНОСТРАННЫМ ЯЗЫКОМ



**В.Ю. Красовский**  
Студент 4 курса БГУИР по специальности "Информатика и технологии программирования"  
bsuiromwi@gmail.com



**С.Н. Нестеренков**  
Декан факультета компьютерных систем и сетей, кандидат технических наук, доцент  
s.nesterenkov@bsuir.by



**И.Г. Скиба**  
Ведущий инженер-программист отдела информационных технологий, ассистент кафедр ЭВС и ЭВМ  
i.skiba@bsuir.by

### **В. Ю. Красовский**

Студент 4 курса БГУИР по специальности "Информатика и технологии программирования".

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцента кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации

### **И.Г. Скиба**

Окончила БГУИР в 2020 году по специальности "Вычислительные машины, системы и сети", магистрант первого года обучения по специальности "Компьютерная инженерия" БГУИР.

**Аннотация.** Развитие технологий и распространение доступности интернета привело к значительным изменениям в мире, и одной из областей, на которые это оказало значительное влияние, является изучение языков. В современном глобализованном мире изучение иностранного языка становится все более важным для общения, образования и карьерного роста. В результате было разработано множество инструментов и методик для изучения и анализа уровня владения иностранным языком.

Целью данного тезиса является обзор различных инструментов и методов анализа уровня владения иностранным языком, их преимуществ и недостатков, а также их полезности в изучении языка и преподавании.

**Ключевые слова:** Владение языком, иностранный язык, обучение, языковые навыки, инструменты, основанные на технологиях, приложения для изучения языков, персонализированное обучение

### **Введение.**

Изучение языка — это процесс, который включает в себя овладение новым языком или повышение уровня владения им. Изучение иностранных языков всегда было сложной задачей. Это требует обширной практики, самоотдачи и много времени. С развитием технологий было разработано несколько инструментов и методов, помогающих студентам и преподавателям оценивать уровень владения иностранным языком. Эти инструменты варьируются от традиционных методов, таких как грамматические упражнения, до современных инструментов, таких как компьютерные программы и приложения.

### **Актуальность.**

Повышение значимости владения иностранными языками привело к появлению разнообразных инструментальных средств, предназначенных для изучения и анализа уровня владения иностранным языком. В работах [3] и [5] показана важность

использование технологий в процессе изучения иностранного языка. Среди таких инструментальных средств особую популярность имеют веб-приложения, такие как Duolingo, Babbel и Rosetta Stone. Они имеют такие функции, как интерактивные уроки, викторины и игры, которые делают обучение более увлекательным и веселым. Эти приложения показали свою эффективность в улучшении процесса изучения языка и набрали популярность среди изучающих языки.

### **Обзор функций инструментальных средств для изучения иностранного языка.**

В работах [1–3] был проведен анализ методов при изучении иностранного языка. Наиболее распространенными функциями инструментальных средств для изучения иностранного языка являются интерактивные уроки, викторины, социальное взаимодействие, адаптивные алгоритмы и функции геймификации процесса обучения. Интерактивные уроки разрабатываются с учетом различных стилей обучения и уровней владения языком. Тесты дадут учащимся возможность проверить свое понимание уроков. Функции социального взаимодействия позволят учащимся общаться с другими учащимися и носителями языка, который они изучают. Как показано в работе [5] адаптивные алгоритмы играют значительную роль в процессе обучения. Они могут регулировать уровень сложности в зависимости от успеваемости учащихся, гарантируя, что они будут испытывать трудности, но не перегружены. Функции геймификации, такие, как баллы, значки и таблицы лидеров, могут мотивировать учащихся участвовать в процессе обучения и отслеживать их прогресс.

Одними из важнейших аспектов инструментальных средств являются их дизайн и пользовательский интерфейс. Дизайн и пользовательский интерфейс веб-приложения для изучения языка должны быть интуитивно понятными и удобными для пользователя. Цветовая гамма и макет должны быть эстетически приятными, а навигация - простой, чтобы учащиеся могли легко найти необходимый им контент. Приложение должно быть оптимизировано как для настольных компьютеров, так и для мобильных устройств, чтобы учащиеся могли получить к нему доступ из любого места.

Интерактивные уроки рекомендуется разрабатывать с учетом различных стилей обучения и уровней владения языком. Уроки должны быть интерактивными, увлекательными и веселыми, чтобы гарантировать, что учащиеся останутся мотивированными и заинтересованными. Уроки, включающие аудио, видео и интерактивные элементы, предоставляют учащимся всесторонний опыт обучения.

Анализ уровня владения иностранным языком обычно представляют собой викторины. Одним из важнейших факторов при разработке викторин, является адаптивный уровень сложности. Викторины должны быть составлены таким образом, чтобы они были достаточно сложными, чтобы стимулировать улучшение уровня владения языком, но не слишком сложными, что может привести к потере мотивации и снижению эффективности обучения. Также важными свойствами викторин являются интерактивность, увлекательность и наличие обратной связи. Викторины рекомендуется разрабатывать с учетом возможности пересдачи, чтобы не навлекать излишний стресс на обучающихся.

Функция социального взаимодействия имеют значимую роль, как показано в работе [4]. Она позволяет учащимся общаться с другими учащимися и носителями языка, который они изучают. Эта функция позволяет учащимся практиковать свои языковые навыки в реальных условиях, а также предоставит им возможность задавать вопросы и получать обратную связь. Функция социального взаимодействия может быть реализована в виде чат-ботов и виртуальных помощников, которые помогут учащимся в изучении языка.

### **Заключение.**

Изучение иностранных языков имеет решающее значение в современном глобализованном мире. Инструменты и методы, используемые для анализа уровня владения иностранным языком, постоянно развиваются, и важно идти в ногу с этими

изменениями. В то время как традиционные методы по-прежнему широко используются, инструментальные средства, такие как веб-приложения для изучения языков, набирают популярность. Использование этих инструментов может помочь учащимся более эффективно развивать свои коммуникативные навыки. Однако эти инструменты не могут заменить роль учителя в предоставлении персонализированной обратной связи и рекомендаций. Будущее изучения языков, вероятно, будет связано с сочетанием традиционных методов и инструментов, основанных на технологиях

#### **Список литературы**

[1] Нестеренков, С.Н. Модель построения расписания на основе прецедентов / С.Н. Нестеренков // Информатизация образования, 2015. - N1. – С. 61-73.

[2] Нестеренков, С.Н. Функциональная модель процедур планирования и управления образовательным процессом как основа построения информационной среды учреждения высшего образования / С.Н. Нестеренков, Н.В. Лапицкая // Вести Института современных знаний, 2018. - N1. – С. 97-105.

[3] Ahmadi, M. R. The Use of Technology in English Language Learning: A Literature Review. / M.R. Ahmadi // IJREE, 2018. – P.116-125

[4] Chun, D. Technology in Language Use, Language Teaching, and Language Learning. / D. Chun, R. Kern, B. Smith // The Modern Language Journal, 2016 – P.64-80

[5] Nor, N. M. A review of theoretical perspectives on language learning and acquisition. / N.M. Nor, R.A. Rashid // Kasetsart Journal of Social Sciences, 2018 – P.161-167

### **DEVELOPMENT OF TOOLS FOR STUDYING AND ANALYZING THE LEVEL OF FOREIGN LANGUAGE PROFICIENCY**

***V.I. Krasovskii***

*4th year student of BSUIR,  
specialty "Computer Science  
and Programming  
technologies"*

***S.N. Nesterenkov***

*Dean of the Faculty of  
Computer Systems and  
Networks of BSUIR,  
PhD of Technical Sciences,  
Associate Professor*

***I.G. Skiba***

*Leading Software Engineer of  
the Information Technology  
Department, Assistant of the  
Department of Electronic  
Computing*

*Department of Computer Science*

*Faculty of Computer Systems and Networks*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus*

*E-mail: bsuiromwi@gmail.com*

**Abstract.** In today's globalized world, learning a foreign language has become increasingly important for communication, education, and career advancement. As a result, many tools and techniques have been developed for studying and analyzing the level of foreign language proficiency.

This thesis aims to review the various tools and methods for analyzing foreign language proficiency, their advantages and disadvantages, and their usefulness in language learning and teaching.

**Keywords:** Language proficiency, foreign language, learning, language skills, technology-based tools, language learning apps, personalized learning

УДК 004.021:004.75

## СТЕГАНОГРАФИЧЕСКИЕ МЕТОДЫ СОКРЫТИЯ ИНФОРМАЦИИ



**И.В. Лосик**

Студент факультета  
компьютерных систем и сети  
по специальности  
“Вычислительные машины  
системы и сети”  
losikivan2002@gmail.com



**С.Н. Нестеренков**

Декан факультета  
компьютерных систем и  
сетей Белорусского  
государственного  
университета  
информатики и  
радиоэлектроники, доцент  
кафедры программного  
обеспечения  
информационных  
технологий  
s.nesterenkov@bsuir.by



**Д.В. Низовцов**

Начальник отдела  
информационных технологий,  
ассистент кафедры ПОИТ  
d.nizovtsov@bsuir.by

### **И.В. Лосик**

Окончил Минский радиотехнический институт. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

### **С.Н. Нестеренков**

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации

### **Д.В. Низовцов**

Окончил БГУИР в 2016 году по специальности "Программное обеспечение информационных технологий", магистрант первого года обучения по специальности "Электронные системы и технологии" БГУИР.

**Аннотация.** В данной работе были рассмотрены принципы стеганографических методов, различные алгоритмы сокрытия информации и инструменты для их обнаружения. Также была поднята тема актуальности развития методов стеганографии и инструментов для их обнаружения в сфере кибербезопасности. Понимание стеганографических методов и их применение может помочь защитить данные от несанкционированного доступа и сохранить конфиденциальность личной информации.

**Ключевые слова:** Стеганография, LSB, криптография, кибербезопасность.

### **Введение.**

Стеганография - это технология, которая позволяет скрыто передавать сообщение в другом сообщении таким образом, чтобы даже его наличие оставалось незамеченным. Эта технология может использоваться для защиты конфиденциальной информации при ее передаче через интернет или другие открытые коммуникационные среды. Стеганографические методы сокрытия информации могут быть использованы в различных областях, например, в криминалистике, государственной безопасности, бизнесе. Они



позволяют скрыть конфиденциальную информацию, например, финансовые данные или персональные данные.

#### **Актуальность.**

Эта технология становится все более популярной в различных областях, в том числе в области кибербезопасности, обмена личными данными, медицинских записей и т.д. Она также может использоваться для целей, связанных с правоохранительными органами и военными. Из-за своей актуальности и важности стеганографические методы сокрытия информации становятся объектом внимания исследований и разработок в области информационных технологий.

#### **Анализ принципов работы метода LSB.**

Одним из примеров стеганографического метода является сокрытие информации в видеороликах или изображениях. Это может быть достигнуто путем изменения пикселей изображения или встраивания информации в частотный спектр аудио- или видеофайла. Другим примером может быть сокрытие информации в текстовых файлах, где каждый символ в файле может использоваться для передачи битов информации.

Далее будет рассмотрен Метод Least Significant Bit.

(LSB) Метод Least Significant Bit (LSB) - это один из самых простых методов стеганографии, который использует наименее значимый бит в каждом байте изображения для сокрытия данных. Этот метод позволяет встраивать данные в любой формат изображения без сильного визуального искажения оригинального изображения. Представим изображение на рисунке 1 в виде матрицы (рисунок 2), в которой каждая ячейка выражает пиксель, а её значение – цвет этого пикселя/ячейки.



– Рисунок 1. Исходное изображение

$$P_{3,3} = \begin{pmatrix} 10000110 & 00101111 & 00000000 \\ 11111111 & 11111111 & 00000000 \\ 11000001 & 00101111 & 00000000 \end{pmatrix}$$

–Рисунок 2. Матрица, составленная на основе изображения

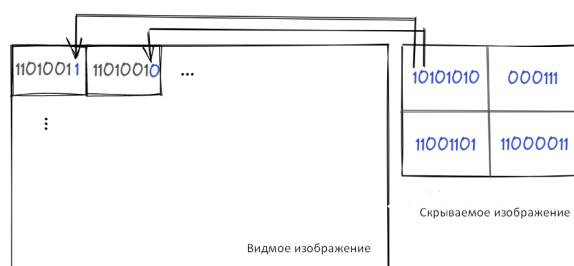
–Изменим младший бит. Для этого реверсируем их и посмотрим, как в итоге будут выглядеть матрица и новое фото (рисунок 3).

$$P_{3,3} = \begin{pmatrix} 10000111 & 00101110 & 00000001 \\ 11111110 & 11111110 & 00000001 \\ 11000000 & 00101110 & 00000001 \end{pmatrix}$$



– Рисунок 3. Матрица и изображение, полученные после изменения младших битов исходного изображения

–Значительного визуального изменения не произошло, но матрица изображения была изменена и именно на этом и построен метод LSB. На рисунке 4 представлен общий принцип этого метода.



– Рисунок 4. Исходное изображение

–Существует еще несколько методов стеганографического сокрытия информации, которые могут быть использованы для передачи информации незаметно. Рассмотрим наиболее распространенные из них:

–1. Метод расширения спектра - этот метод используется для сокрытия информации в частотном спектре звукового или видео файла. Для этого используются математические алгоритмы, которые изменяют соотношение между высокими и низкими частотами в звуковой или видеодорожке. Однако этот метод может изменять качество звука или видео, что может привести к потере информации.

–2. Метод сокрытия в фоне - этот метод использует несколько изображений или фрагментов изображений, чтобы скрыть информацию. Например, можно скрыть сообщение в градиентах цвета или в малозаметном участке изображения. Этот метод является достаточно сложным и менее эффективным, но при правильном использовании может быть достаточно эффективным.

–3. Метод сокрытия в текстовых файлах - этот метод использует изменение символов в текстовых файлах для сокрытия информации. Например, каждый символ может быть использован для передачи одного бита информации. Этот метод является одним из самых простых и эффективных, но может быть детектирован специальными программами, что может привести к раскрытию информации.

–4. Метод сокрытия в сетевом трафике - этот метод используется для передачи информации через сетевой трафик. Например, информация может быть скрыта в интервалах между пакетами или в качестве фонового шума. Этот метод также достаточно сложен и может быть детектирован специальными программами, что может привести к раскрытию информации.

### **Заключение.**

В целом, можно сказать, что стеганографические методы сокрытия информации являются важным инструментом для защиты конфиденциальных данных в различных сферах жизни. Они позволяют закодировать и скрыть информацию внутри другой информации, что делает ее неразличимой при поверхностном рассмотрении. Важно отметить, что развитие стеганографических методов и инструментов для их обнаружения является актуальной темой исследований в области информационных технологий и кибербезопасности. Необходимо прикладывать все усилия для того, чтобы гарантировать безопасность и конфиденциальность обработки и передачи информации.

### **Список литературы**

- [1] Грибунин, Вадим Геннадьевич Стеганография, цифровые водные знаки и стеганоанализ / Грибунин Вадим Геннадьевич. - М.: Вузовская книга, 2018. - 110 с.
- [2] Земцов, Андрей Методы цифровой стеганографии для защиты авторских прав: моногр. / Андрей Земцов. - М.: LAP Lambert Academic Publishing, 2019. - 148 с.
- [3] Рассел, Джесси Стеганография / Джесси Рассел. - М.: VSD, 2019. - 193 с.
- [4] Рябко, Б. Я. Основы современной криптографии и стеганографии / Б.Я. Рябко, А.Н. Фионов. - М.: Горячая линия - Телеком, 2019. - 232 с.
- [5] Рябко, Борис Яковлевич Основы современной криптографии и стеганографии / Рябко Борис Яковлевич. - М.: Горячая линия - Телеком, 2019. - 166 с.
- [6] Орлов Д. В., Нестеренков, С. Н., Марков А.Н., Сравнительный анализ методов бинаризации изображений [Электронный ресурс]URL:<https://libeldoc.bsuir.by/handle/123456789/4834>
- [7] Ситников А.В., Нестеренков С.Н., Марков А.Н. Методы защиты больших данных [Электронный ресурс]. URL: <https://libeldoc.bsuir.by/handle/123456789/48384>.

## **QUALITY EVALUATION OF INFORMATION TRANSFER IN A DISPATCHING SYSTEM BASED ON MQTT ARCHITECTURE**

***I.V. Losik***

*Student of the Faculty of  
Computer Systems and  
Networks*

***S.N. Nesterenkov***

*Dean of the Faculty of  
Computer Systems and  
Networks of the Belarusian  
State University of  
Informatics and  
Radioelectronics, Associate  
Professor of the Department  
of Software information  
Technology*

***D.V. Nizovtsov***

*Head of Information  
Technology  
Department, assistant of the  
department of Software  
information Technology*

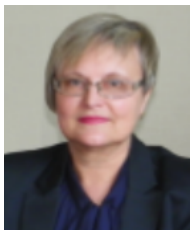
*Department of Information and Computer Systems Design  
Faculty of Computer Systems and Networks  
Belarusian State University of computer science and Radio Electronics, Republic of Belarus  
E-mail: alexvikt.minsk@gmail.com*

**Annotation.** In this paper, the principles of steganographic methods, various algorithms for hiding information and tools for their detection were considered. The topic of the relevance of the development of steganography methods and tools for their detection in the field of cybersecurity was also raised. Understanding steganographic techniques and applying them can help protect data from unauthorized access and keep personal information private.

**Keywords:** Steganography, LSB, cryptography, cybersecurity.

УДК 339.38

## COMPOSABLE BUSINESS IN DIGITALTRANSFORMATION



**L.I. Arkhipova**

*Associate professor, PhD, BSUIR*

### **L.I. Arkhipova**

*Graduated from the Belarusian State University and Academy of public administration under the aegis of the President of the Republic of Belarus. Has years of experience in Semiconductor industry (Integral, Minsk). Conduct scientific researches in the field of digital business transformation and marketing technologies.*

**Abstracts.** This article is devoted to the new trend in digital business and new types of activity, which has a name «composable business». Composable business essence explained in the article and described how this business works. Besides, the possible ways of a business modularization and reconfigurations presented in this article to satisfy the needs of the current situation and concrete customer requirements. Composable business realization go through the composable technologies and composable thinking to reach business benefits.

**Key words:** composable business, composable thinking, composability, modularity, autonomy, discovery, orchestration, enterprise architecture, principles of composability, agile business.

### **Introduction.**

Gartner's experts determine composable business as follows – business that can be easily modularized and reconfigured to meet the needs of the current situation. This can be done through composable technologies, which allow different parts of the business to be connected and worked together. That will enable businesses to quickly adapt to changes in the market and respond more effectively to customer needs [1-3].

«Composable business is a natural acceleration of digital business that you live every day. It allows us to deliver the resilience and agility these interesting times demand» (Daryl Plummer – Distinguished VP Analyst, Gartner) [1]. In addition, specialists emphasis – composable business allows quickly adapt activity to market changes by assembling the right combination of applications and services for each situation. Composable business is primarily a mental concept and secondarily a technology. This change of order actually reflects the whole evolution in digital era [2-4]. Using composable technologies, businesses can create independent modules that can still work together when needed. That gives enterprises the flexibility they need to respond quickly to changing business conditions.

**Research and discussion.** Currently, composable business defined as *intelligent composable business* because it improves and transforms decision-making on the base of actual structured data by accessing on Big Data and Artificial intelligent methods of analysis. Intelligent composable business deals with the following activities: generate new business models; support autonomous operations; develop new products/services through prototyping and MVP; develop and support omnichannels system, use CX and EX as business assets. In a nearest future, it will become increasingly important to create agile and adaptive systems that can quickly respond to changing market conditions.

Definitions of composable business can be determined through the term «composability».

*Composability* – software design approach where individual components arranged, re-arranged, and discarded when need to use a no-code platform.

*Business composability* – technology under which businesses become more agile by replacing big quantity of applications and data with composable building blocks that are *modular*

and interchangeable. From this point of view, composability implies a *new operating model* that involve and affect *people and processes*. Therefore, organizations need to consider and make the decision about basic building blocks of a composable business [2-5]:

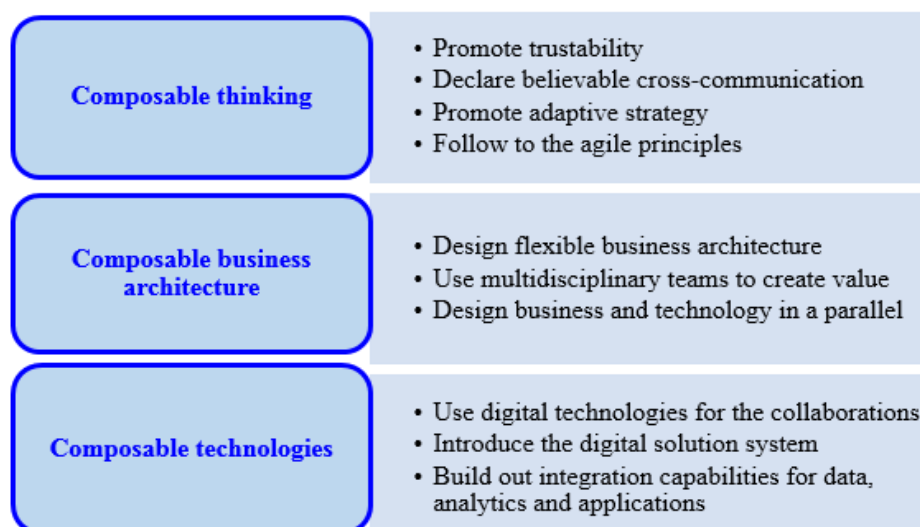
*Composable thinking (People)*: focusing on adaptive strategy, flexible organizational structures, and planning cycles. Composable thinking determines the ability to think in modular terms and recognize the business as a collection of independent but interconnected parts and cycles. Composable thinking supports the flexibly – conduct quick changes due to market needs change. Main idea in realization of this block to define the concept – what to compose and when. Such approach practically changes business capabilities and market opportunity.

*Composable business architecture (Processes)*: focusing on value-based processes, working with business capabilities, distributed accountability. Composable thinking does not work without a composable architecture, which represent the different business elements –capabilities, products, teams, processes, services, etc. Such approach dynamically evolves to create new value. *Engagement, technology, and capability* are three dimensions of composable business architecture, which support business growth and scalability.

*Composable technologies (Technology)*: characterized by agile development methodologies (Lean Agile principles – SAFe framework), modular architecture, and distributed data processing. Composable thinking and architecture with the right technology assets and capabilities practically promote the usage of modular components and their reusability. Such approach supports automation in marketing in business.

*Composable technologies* – are the tools or software, that include *no-code* and traditional *BI platforms*, that enable businesses to modularize their operations and connect different parts of the business. Such approach improves agility and flexibility to enhance the Business-to-Business customer experience (CX).

Interaction blocks, teams, software and customers support the competitive advantages (picture 1).



Picture 1. Composability principles in action [2-5]

In order to introduce *new operating model on the base of composable business concept* need to do the followings:

–Prepare company to the introduction the composable architecture – assign champion and create agile centers of expertize and agile coaches with a mission to embed composability principles into enterprise system.

–Estimate an enterprise position (composability index) and describe the essence of composability principles.

–Specify the skills of employees, starting from methodologies (agile and scrum) until architecture (including microservices) and business expertize.

–Choose the platform that allows to connect enterprise core system, and that supports *integration strategy*, which based on the three composable pillars: thinking, architecture, technology.

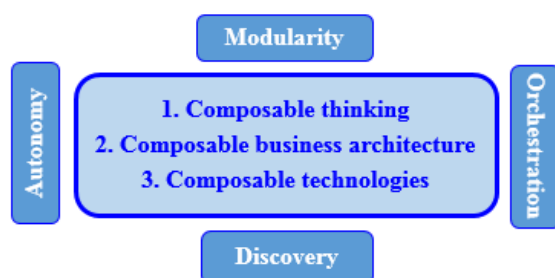
–Choose and implement building unit of business – *Packaged Business Capability (PBC)* (as a Lego bricks). Each PBC implements a business-relevant function – together they create a *digital solution*. PBCs defined as a combination of *microservices* that form independent business units. In this context, *microservices* are small, independent services that work together to reach business goals. A single PBC could provide a complete application implementation to end customers (it is usually a set of business capabilities – PBCs).

Practical realization of this concept based on a combination of such principles as: discovery, modularity, autonomy, orchestration (table 1) [2-6].

Table 1. Four Key Principles of Composability

Principles	Essence	Core meaning
Discovery	More speed through discovery	When you need to move quickly, you need to be able to discover the right combination of applications and services to fit your current needs.
Modularity	Greater agility through modularity	Modularity allows businesses to change and update parts of their applications without affecting the rest of the system. That makes it easier to respond to market or customer needs changes.
Orchestration	Better leadership through orchestration	Orchestration allows businesses to manage different parts of the business. This gives businesses greater control over their operations and helps them to respond more effectively to changes. Orchestration focuses on security components, API support workflow orchestration.
Autonomy	Resilience through autonomy	Autonomy allows businesses to continue operating even if one part of the system fails. That helps businesses to avoid disruptions and keep their operations running smoothly

Balanced combination of the compacibility principles around key blocks is a shortage way to transfer them into competitive advantages in digital economy (picture 2).



Picture 2. Design Principles of Composability (adapted) [3-5]

The value of composability for business could be explain through *business perspective and development perspective* [6].

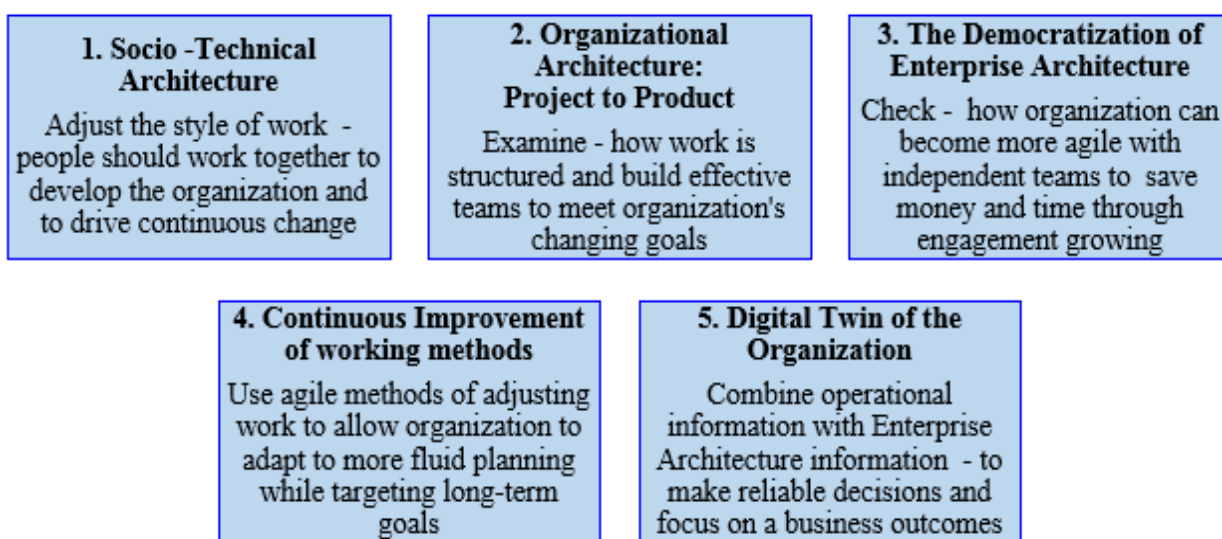
*Business perspective* – business composability gives organizations the agility to react and respond to market changes and adapt faster to the customer requirements [2]. The modular composition of applications allows for bridging the traditional gap between business and IT since both business and users can assist the tasks needed to assemble the experience to the «job to be done» (JTBD). Combination of applications and experiences allow going to market faster.

*Development perspective* – business composability increases developer productivity exponentially because developers start the projects with ready building blocks.

Benefits of business composability expressed as a followings:

1. Composability allows businesses to be more agile and responsive to change, which is essential in today’s fast world.
2. Composability allows businesses to become more efficient and resilient and to manage their operations better.
3. Composability makes it easier for businesses to connect different parts of their operation, improving productivity and communication.
4. Building blocks of application can be assembled quickly and easily, allowing organizations to get new products and services to market faster.
5. Composable systems are more reliable as faulty components can be replaced without affecting the entire system.
6. The right combination of assembled applications and services can optimize businesses operations and resources. This can help reduce costs and improve performance.
7. Composability destroys limitations of monolithic applications that do not support flexibility or customization.

Specialist in business declare that only CHANGE is certain in business environment today. So, need to investigate and follow to the enterprise architecture trends in order to align main business activities and initiatives with business transformation strategies. To understand the top strategic enterprise architecture trends – means to become more agile and deliver change and innovation faster, offer greater value to the market, and meet the organizational objectives (picture 3) [6-9].



Picture 3. Enterprise Architecture Trends (adapted)



## **Conclusions and recommendations.**

A composable business model – is an acceleration model of the digital era. Designing business processes around this idea, forces businesses to be resilient and creates a culture of stability within digital environments. Very important to understand – how organizations apply new concepts into digital business execution. Intelligent composable business improves and even transforms decision-making on the base of actual structured data by accessing on Big Data and Artificial intelligent flexible method of analysis. Composable business will allow create *new business models*, support autonomous operations, develop new products, services, develop and support channels, use CX and EX as a business assets.

## **Reference**

[1] What is Composable Business? (Definition, Principles & Benefits) [Electronic resource]. - An access Mode: <https://www.erp-information.com/composable-business> (Date of access: 12.01.23).

[2] The Core Design Principles of Composability, Becoming Composable: A Gartner Trend Insight, Gartner, 17 Sept 2021.

[3] Becoming a Composable Business [Electronic resource]. - An access Mode: <https://www.outsystems.com/blog/posts/composable-business/> (Date of access: 24.02.23).

[4] What is a composable enterprise? And who is the composer? [Electronic resource]. - An access Mode: <https://blogs.sap.com/2022/12/22/what-is-a-composable-enterprise-and-who-is-the-composer/>

[5] Composable Digital Heart. What is Composable Business? [Electronic resource]. - An access Mode: <https://www.sofokus.com/what-is-composable-business/> (Date of access: 18.02.23).

[6] Архипова, Л.И. Компонуемый бизнес как тренд в цифровой ТРАНСФОРМАЦИИ//Л.И. Актуальные вопросы фундаментальных и прикладных научных исследований // Сборник научных статей по материалам I Международной научно-практической конференции (24 января 2023 г., г. Уфа) / – Уфа: Изд. НИЦ Вестник науки, 2023. – С.58-64 (Date of access: 16.04.23).

[7] What is Intelligent Composable Business? <https://smartz-solutions.com/what-is-intelligent-composable-business/> (Date of access: 23.03.23).

[8] Building Blocks of Modeling a Composable Business ) [Electronic resource]. - An access Mode: <https://www.ardoq.com/blog/composable-business> (Date of access: 12.04.23).

[9] The Top 5 Strategic Enterprise Architecture Trends [Electronic resource]. - An access Mode: <https://content.ardoq.com/enterprise-architecture-trends?hsCtaTracking=7e4c3cb0-2353-4668-8884-b8500e582b5e%7C12a9184a-a7e8-4a36-bae4-961a81780fef> (Date of access: 10.04.23).

## **КОМПОНУЕМЫЙ БИЗНЕС В УСЛОВИЯХ ЦИФРОВОЙ ТРАНСФОРМАЦИИ**

***Л.И. Архипова***

*Доцент кафедры экономики Белорусского  
государственного университета  
информатики и радиоэлектроники,  
кандидат экономических наук, доцент*

*Учреждение образования Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь.*

*E-mail: l.arkhipova@gmail.com*

**Аннотация.** Данная статья посвящена новому тренду в цифровом бизнесе и различным видам активностей, которые названы компонуемым бизнесом. Определена сущность компонуемого бизнеса и описано, как этот бизнес работает. Кроме того, в статье представлены возможности перехода на модульность и реконфигурацию бизнеса с целью удовлетворения требований клиентов в условиях изменения текущей ситуации. Для достижения рыночных выгод компонуемый бизнес реализуется на базе компонуемых технологий и мышления, а также базируется на компонуемой бизнес-архитектуре.

**Ключевые слова:** компонуемый бизнес, компонуемое мышление, компонуемая бизнес- архитектура, компонуемость, модульность, автономность, оркестровка, принципы компонуемости, гибкий бизнес.



UDC 616-009.5; 004.934

## PRESENTATION AND PROCESSING OF DATA FOR THE DIAGNOSTICS OF NEUROLOGICAL DISEASES



**V.A. Vishnyakov**

*Doctor of Technical Sciences,  
Professor of the Department of  
Infocommunication Technologies,  
BSUIR  
vish@bsuir.by*



**Yu ChuYue**

*PhD student in the Department  
of infocommunication  
technologies at BSUIR.  
ycy18779415340@gmail.com*



**Xia YiWei**

*PhD student in the  
Department of  
infocommunication  
technologies at BSUIR.  
903016812@qq.com*

### **V.A. Vishnyakov**

*Area of professional interests/research: information management and security in infocommunications, Internet of things, blockchain, IT in education, intelligent control systems.*

### **Yu ChuYue**

*PhD student in the Department of Infocommunication technologies of the BSUIR.*

### **Xia YiWei**

*PhD student in the Department of Infocommunication technologies of the BSUIR.*

**Annotation.** Big data is triggering a revolution in scientific research mindset. The commonality of data, the overall characteristics of the network are hidden in the data network. In the era of big data, IT systems need to change from having data around the processor to processing power revolves around the data. The processing and presentation of data becomes one of the significant evaluation criteria to measure the effectiveness of the model in performing the task. This article focuses on the data processing process and presentation for diagnosing neurological diseases. Based on the experiments, the data is described, presented and processed.

**Key words:** Machine learning, data processing, data description, data presentation.

### **Introduction.**

With the rapid development of global information science and technology over the Internet, the term "big data" is increasingly familiar to people [1].

In a general sense, Big Data is a collection of data that cannot be sensed, accessed, managed, processed, and served within a tolerable period of time using traditional IT technologies as well as software or hardware tools. In 2008, «Nature» published a specific journal "Big Data" [2], which introduced the challenges posed by massive data from various aspects. A special edition of «Science» on data processing,

"Dealing with data" [3], was launched in 2011 to discuss the challenges posed by the Data Deluge. The aim is to discuss big data applications in biomedicine, Internet, economics, environmental science, supercomputing, etc.

The high integration of the triadic worlds of human, machine, and material has triggered an explosion in the scale of data and a high degree of complexity in data patterns.

Traditional data-centric disciplines, such as medicine, are generating more and more data, which hide great opportunities and values.

Based on the source of data, big data can be divided into two broad categories in a preliminary way: a category from the physical world and another from human society. The former is mainly scientific experimental data or sensor data, while the latter is related to human activities,

especially associated with the Internet [4].

These two types of data vary tremendously in the way they are processed and their targets, with different ways of presentation of the data.

In a general sense, data used to diagnose neurological diseases usually include various types of clinical data, imaging data, physiological data, and molecular data [5], which can help physicians gain a more comprehensive understanding of patients' conditions, perform diagnosis, therapy, and prognosis assessment of neurological diseases, etc., it can also be used to train machine learning models such as neural networks to enable automated disease diagnosis and prediction.

In this article, we took Alzheimer's disease and Parkinson's disease as examples, specifically, the way voice data used to diagnose neurological diseases are processed and presented in machine learning models.

#### **Data description.**

Parkinson's disease and Alzheimer's disease both have their own public datasets due to being internationally recognized as incurable neurological diseases. Among the voice datasets available in the field of Parkinson's disease diagnosis are the UCI Machine Learning Library [6], the mPower dataset [7], and the Parkinson's Disease Data Set [8].

In the case of the UCI Machine Learning Library, the data used in this dataset was collected from 188 PD patients (107 males and 81 females), aged between 33 and 87 years old, as shown in Figure 1.

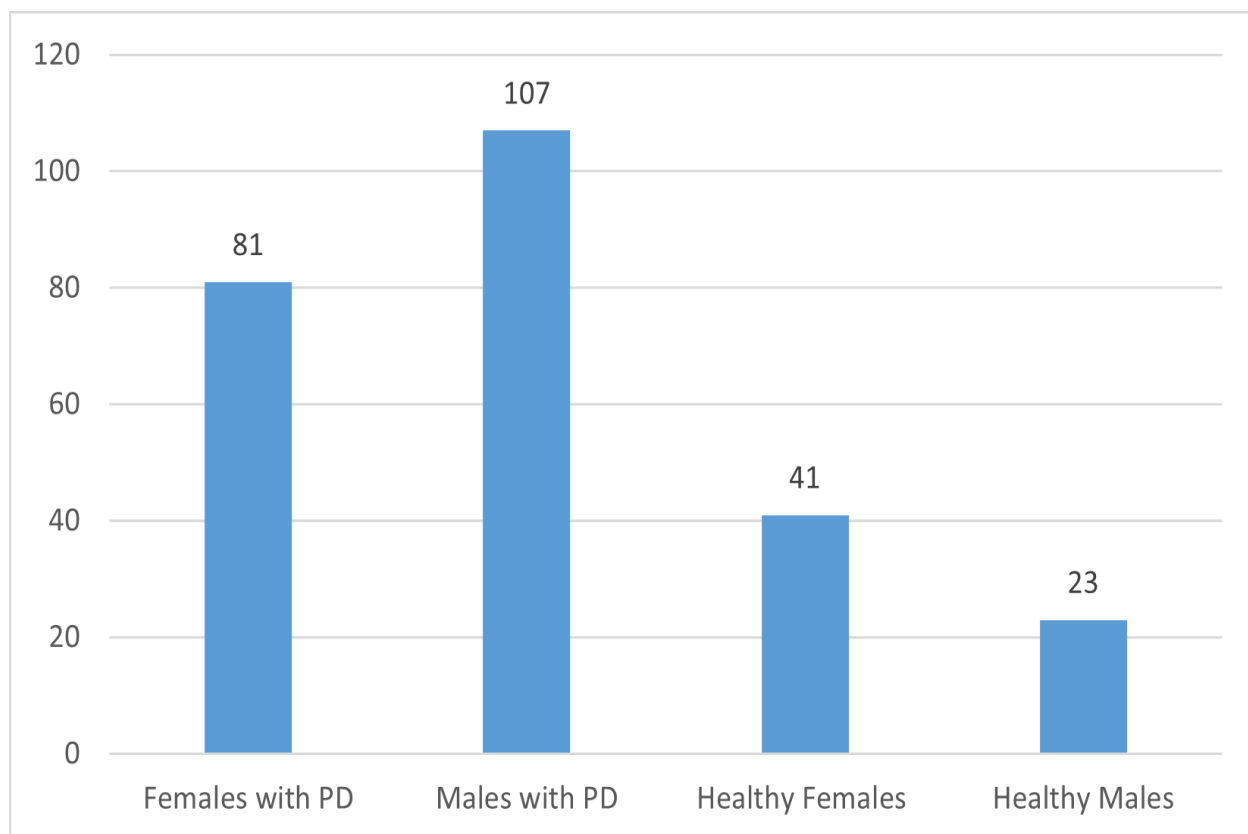


Figure 1. Data composition of the UCI Parkinson's dataset

The control group consisted of 64 healthy individuals (23 males and 41 females), aged between 41 and 82 years old. During data collection, the microphone was set to 44.1 KHz with

sustained pronunciation of the vowel /a/ collected from each participant after a physician's examination, repeated three times, and the data was audio format with a .WAC suffix, each sample containing 754 attributes, of which 752 attributes were phonetic features, 1 attribute identified age, and 1 attribute identified gender type.

In the field of Alzheimer's disease diagnosis, there are voice datasets such as DementiaBank [9], Dem@Care Datasets [10], and The ADReSS Challenge [11], a competition dedicated to providing a spontaneous voice benchmark dataset to test the effectiveness of different methodologies for voice recognition in Alzheimer's disease. In the case of The ADReSS Challenge 2020, for example, in order to minimize the risk of bias in the prediction task, the data used in this dataset was collected from 2,122 voice fragments of 78 non-affected individuals and 2,155 voice fragments of 78 patients in audio format with a .WAC suffix, with each voice fragment associated with a corresponding transcript, the participants were in the age range of 50-80 and supported both phonological and semantic analysis of the data.

### **Data presentation.**

The presentation of Parkinson's voice data and Alzheimer's voice data in machine learning systems usually requires multiple steps, including data preparation, data training and data evaluation, etc. The visualization of data helps users understand and analyze the data, as illustrated below:

#### **A. Data preparation.**

Data preparation includes the steps of data collection, organization, processing and filtering with the aim of ensuring the accuracy and usability of data. The concrete content of data collation is to remove or recover missing values, duplicate data, outliers, erroneous data, etc.

#### **B. Data training**

In order for machine learning models to better recognize the data, it is necessary to convert the data into a format suitable for model training, such as converting text into vectors, extracting features from voice, etc. In the voice assessment of Parkinson's, many of which have been shown to have varying degrees of outliers compared to healthy populations, the most commonly used acoustic characteristics include: fundamental frequency (F0) and fundamental frequency variation (vF0), frequency perturbation (Jitter), amplitude perturbation (Shimmer), and harmonic-to-noise ratio (HNR).

In addition, other acoustic metrics such as MFCC (Mel Frequency Cepstral Coefficients), recurrence period density entropy, detrended fluctuation analysis (DFA), and VOT (Voice Onset Time) have also gained importance. In the voice assessment of Alzheimer's disease, the variables of diagnostic tests are usually defined according to preclinical symptomatology related to semantic memory and verbal fluency tasks.

In addition, algorithms such as CVT (Cepstral Vowel Transform) algorithm, or acoustic feature sets such as eGeMAPS, ComParE, emobase, etc., or feature extraction tools such as YAAFE (Yet Another Audio Feature Extractor), COVAREP (ComParable Voice Analysis & REpresentation Patterns), OpenSMILE toolkit can be used to help improve the accuracy of the model.

If the features are too redundant or to optimize the network parameters, an algorithm is considered for feature selection, where the most relevant or representative features are selected for the voice data to improve the efficiency and accuracy of the model.

#### **C. Data evaluation.**

During the training process, visualization tools can be used to monitor the training effect of the neural network, such as plotting loss function images and making accuracy variation tables, all of which can help to tune the parameters and optimize the network structure to improve the performance of the model.

After the training is completed, in order to evaluate the model, calculating metrics such as accuracy, recall, F1 score helps to compare with the effect of other models or visualize the results using confusion matrix and ROC curve to visually show the accuracy of the network.

### Data processing.

In machine models, data presentation and processing are usually closely related steps, except that data processing focuses more on processing, extracting, and preprocessing the data to facilitate subsequent analysis, while data presentation focuses more on the mode and content of presenting the data.

The presentation of data can directly affect the subsequent data processing, and the results of data processing can in turn affect the way the data are presented.

Figure 2 shows how the Parkinson's voice data was processed in reference [12].

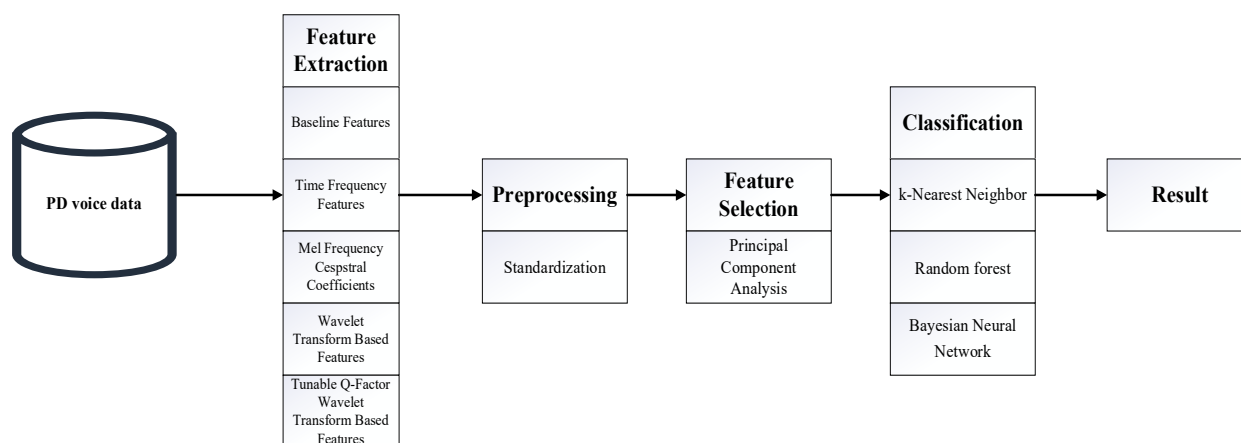


Figure 2. Processing of Parkinson's data

### Conclusion.

Combining the above information, it can be seen that data used to diagnose neurological diseases usually go through several processing steps in machine learning models.

This report describes the presentation and processing of voice data in machine learning models, taking the data of Alzheimer's disease and Parkinson's disease as examples.

Data plays a crucial part in the diagnosis of neurological diseases, safeguarding the reliability, accuracy and integrity of data is a prerequisite for models to make good judgments.

In general, in the future, one needs to overcome the difficulties of data sparsity and redundancy as well as enhance the ability to obtain knowledge from massive and complex data in order to improve the efficiency of research and production so as to move the IT-diagnostics to a new stage of digitization and informatization.

### References

- [1] Zheng, Yili. Research on Big Data Governance in Public Hospitals – A Case Study of Wenzhou E Hospital. Master's thesis. / Zheng Yili. – Shanghai Normal University. 2021. – 82 c.
- [2] Big Data. Nature. – 2008. – 455(7209). –136 c.
- [3] Dealing with data. Science. – 2011, – 331(6018). – P. 639–806.
- [4] Li Guojie, Cheng. Research Status and Scientific Thinking of Big Data / Li Guojie, Cheng Xueqi. – Journal of the Chinese Academy of Sciences – 2012. – 27(06). – P. 647–657.
- [5] Rui, Zhang. Advancing Alzheimer's research: a review of big data promises / Rui Zhang, Gyorgy Simon, Fang Yu. – Int J Med Inform. 2017. – P. 48–56.
- [6] Sakar, C.O. A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. / Sakar, C.O., Serbes, G., Gunduz, A., Tunc, H.C., Nizam, H., Sakar, B.E., Tutuncu, M., Aydin, T., Isenkul, M.E. and Apaydin, H. – Applied Soft Computing, 2019. – P. 255–263.

[7] Brian, M. Bot. The mPower study, Parkinson disease mobile data collected using ResearchKit. / Brian M. Bot, Christine Suver, Elias Chaibub Neto, Michael Kellen, Arno Klein, Christopher Bare, Megan Doerr, Abhishek Pratap, John Wilbanks, E. Ray Dorsey, Stephen H. Friend and Andrew D. Trister. . Scientific Data 3, 2016. – 9 p.

[8] Max, A. Little. Suitability of dysphonia measurements for telemonitoring of Parkinson's disease / Max A. Little, Patrick E. McSharry, Eric J. Hunter, Jennifer Spielman, Lorraine O. Ramig. – Nature Precedings. 2008. – 27 p.

[9] Alyssa, M. Lanzi. DementiaBank: Theoretical rationale, protocol, and illustrative analyses / Alyssa M. Lanzi, Anna K. Saylor, Davida Fromm, Houjun Liu, Brian MacWhinney, Matthew L. Cohen. – American Journal of Speech-Language Pathology. 2023. – P. 426–438.

[10] Karakostas, A. The Dem@Care Experiments and Datasets: a Technical Report. / Karakostas A., Briassouli A., Avgerinakis K., Kompatsiaris I., Tsolaki M. – ArXiv. 2016. – 4 p.

[11] Saturnino, Luz. Alzheimer's dementia recognition through spontaneous speech: The ADReSS challenge. / Saturnino Luz, Fasih Haider, Sofia de la Fuente, Davida Fromm, Brian MacWhinney. - ArXiv, 2020. – 5 c.

[12] Vishniakou, U. A. IT Diagnostics of Parkinson's Disease Based on the Analysis of Voice Markers and Machine Learning. / Vishniakou U. A., Xia YiWei. – Doklady BSUIR. 2023; 3(23).

## **ПРЕДСТАВЛЕНИЕ И ОБРАБОТКА ДАННЫХ ДЛЯ ДИАГНОСТИКИ НЕВРОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ**

***В.А. Вишняков***

*Профессор кафедры  
инфокоммуникационных  
технологий, доктор технических  
наук*

***Ю. Чуйэ***

*Аспирантка кафедры  
инфокоммуникационны  
х технологий*

***С. Ивей***

*Аспирант кафедры  
инфокоммуникационных  
технологий*

*Кафедра инфокоммуникационных технологий*

*Факультет информационной безопасности*

*Белорусский государственный университет информатики и радиоэлектроники,*

*Республика Беларусь*

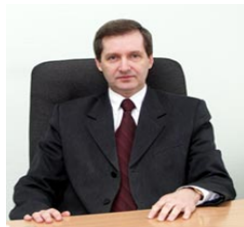
*E-mail: vish@bsuir.by*

**Аннотация.** Обработка и представление данных становится одним из важных направлений машинного обучения при выполнении задач ИТ-диагностики. Доклад посвящен процессу представления, обработки голосовых данных пациентов и их представлению для диагностики неврологических заболеваний. Наборы данных для обучения нейронных сетей описаны, представлены и обработаны.

**Ключевые слова:** машинное обучение, описание голосовых данных, представление данных, обработка данных.

UDC 004.021, 53.087.44, 612.172.4

## HEART RATE MEASUREMENT ALGORITHM IN THE SYSTEM OF REMOTE MONITORING OF HUMAN CONDITION



**A.N. Osipov**

Associate Professor of the  
Department of Electronic  
Engineering and Technology,  
PhD  
osipov@bsuir.by



**O.Ch. Rolich**

Associate Professor of the Department of  
Information and Computer Systems  
Design, PhD  
aoleg.rolich@gmail.com



**M.P. Batura**

Professor, Doctor of  
Sciences



**A.P. Kluev**

Senior Lecturer  
of the Department of  
Engineering Psychology  
and Ergonomics  
kluev@bsuir.by



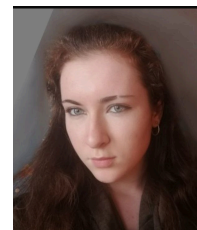
**V.D. Vladymtsev**

Assistant  
of the Department  
Computer Science  
v.vladymtsev@bsuir.by



**S.A. Migalevich**

Head of the Center for  
Informatization and  
Innovative Developments,  
Master of Technical  
Sciences  
migalevich@bsuir.by



**E.R. Uglanova**

Student of BSUIR

### **A.N. Osipov**

Associate Professor of the Department of Electronic Engineering and Technology. Research interests: biomedical engineering.

### **O.Ch. Rolich**

Associate Professor of the Department of Information and Computer Systems Design. Engaged in digital signal processing.

### **M.P. Batura**

Professor, Doctor of Sciences. Research interests: system analysis, information processing in technical and organizational systems.

### **A.P. Kluev**

Senior Lecturer of the Department of Engineering Psychology and Ergonomics. Research interests: biomedical engineering.

### **V.D. Vladymtsev**

Assistant of the Department Computer Science. Research interests: neural network analysis.

**S.A. Migalevich**

*Head of the Center for Informatization and Innovative Developments. Research interests: neural network analysis.*

**E.R. Uglanova**

*Student of BSUIR. Research interests: biomedical engineering.*

**Abstract.** The article discusses the issues of measuring the heart rate during human movement in remote monitoring systems. The algorithm of heart rate measurement based on the processing of spectrograms of cardiograms is described. The implementation of the algorithm makes it possible to obtain a more accurate quantitative assessment of the heart rate in real time in the presence of interference and noise.

**Keywords:** heart rate, spectrogram, electrocardiogram, real time scale.

### **Introduction.**

The explosive growth of digital medicine is facilitated by new solutions in the field of information technology, artificial intelligence, sensors, robotics, wireless communications, processing and analysis of large amounts of information, augmented and virtual reality. The growth of the industry's demand for IT-solutions is also associated with an increase in the proportion of patients with chronic diseases and necessity to provide them with constant monitoring and long-term care. The coronavirus pandemic has become a powerful stimulus for the development of digital medicine. Information technologies are used in all structural elements of digital medicine: to improve the accuracy of diagnosis and personification of treatment based on medical big data collected within the electronic medical document flow; to analyze medical images obtained by a telemedicine system using computer vision algorithms, etc. One of the promising applications of information technology is the determination of the functional state of a person, the identification of pre-crisis and crisis states through wearable medical electronics systems, including systems for continuous monitoring of the condition of patients. The basic (prototypes) remote monitoring systems, that has currently being developed are focused on real-time monitoring of the following basic physiological parameters: heartbeat, body and environment temperature, tissue oxygen saturation, body position [1,2,3]. It is an important and difficult issue understand how to measure heart rate (HR) via electrocardiograms (ECG) in real time against the background of moderate common-mode interference, mainly induced in power circuits. Wireless body area network (WBAN), which is currently being developed by various foreign companies, as a rule, implement the function of heart rate monitoring[4,5]. In portable real-time devices, the dynamics of a person and his internal organs impose additional complexity on interference in power circuits when measuring heart rate and other vital signs.

### **Description of the algorithm.**

In this paper, to solve the problem, it is proposed to measure heart rate in real time by analyzing the spectrogram of the electrocardiosignal [6,7,8]. This decision is due to the quasi-periodicity of the heartbeat process and the ability to control the behavior of the signal in certain frequency ranges, especially in the ranges least affected by interference in power circuits, as well as the pronounced nature of the QRS-complex of the electrocardiogram, which, as a rule, clearly stands out against the background of noise.

Figure 1 ( $U$  is the amplitude of the cardiogram,  $N$  is the sample numbers proportional to the current time of the electrocardiographic signal) shows a fragment of the original electrocardiographic signal with QRS-complexes recorded by a portable electrocardiograph from a person in his natural conditions. In this case, the isoelectric line of the electrocardiogram fluctuates along with its elements. This complicates the process of detection and recognition of the QRS-complex. Human movements additionally noticeably distort the signal, in some places almost completely suppressing the QRS-fragment. Sampling period of signal samples  $t_d = 2$  ms.



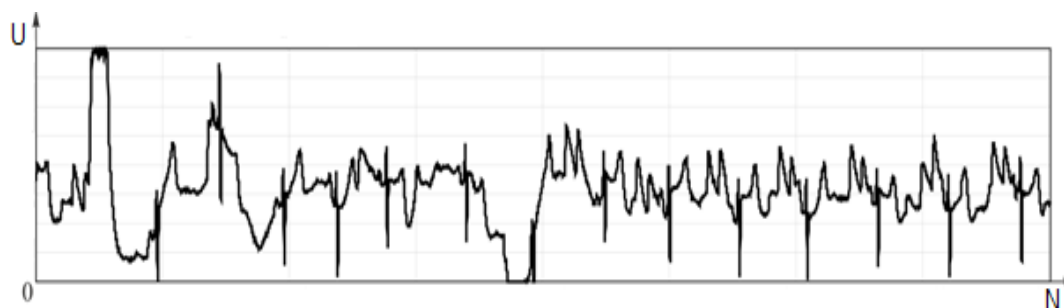


Figure 1. Fragment of the initial electrocardiosignal with QRS-complexes recorded by a portable electrocardiograph

The proposed algorithm for heart rate measurement consists of six stages:

1. ECG registration and calculation of the sliding window fast Fourier transform (FFT) using the Hamming weight function and calculation of the electrocardiogram spectrogram.
2. Formation of the first numerical array, the columns of which are the calculated amplitude spectra of the window Fourier transform, and the lines are frequency slices of the spectrogram on the range  $[f_1, f_k]$ .
3. Calculation of the second Fourier transform of the numeric values of the strings of the first numeric array and the formation of the second numeric array.
4. Search for the coordinates of global highs in the calculated amplitude spectra of the second array.
5. Calculation of heart rate estimates for each amplitude spectrum of the second array, taking into account the weights of the harmonics adjacent to the global highs and the decreasing nature of the fragment of the amplitude spectrum.
6. Averaging of the set of heart rate estimates calculated at the previous stage.

Based on these steps, the heart rate measurement algorithm is based on a two-fold FFT of the initial electrocardiosignal. Heart rate is estimated by repeated FFT of frequency slices of the spectrogram in the range  $[f_1, f_k]$ . A second numerical array is formed, the rows of which are the amplitude spectra obtained as a result of calculating the second Fourier transform. Figure 2 shows one of the amplitude spectra (a sequence of string values of the second numeric array), where  $L$  is the harmonic number of the spectrum proportional to the frequency. Our own research has shown that in the heart rhythm of an electrocardiogram, the spectral region from approximately  $f_1 = 12$  Hz to  $f_k = 45$  Hz is least susceptible to interference. Accordingly, the heart rate measurement in the proposed algorithm is based on peak analysis of the amplitude spectra of the second numerical array, based on the data of frequency slices of the spectrogram with column numbers (or rows, depending on the method of representation of the spectrogram) corresponding to frequencies from the range  $[f_1, f_k]$ .

The decreasing nature of the fragment of the amplitude spectrum of the frequency slice means movement from its maximum in both directions to the first rises. Changes in the amplitude of a certain harmonic component of the original signal there is a quasi-periodic, heartbeat-related process that manifests itself in the spectrogram. Its frequency corresponds to the near-peak frequency response region in Figure 2.



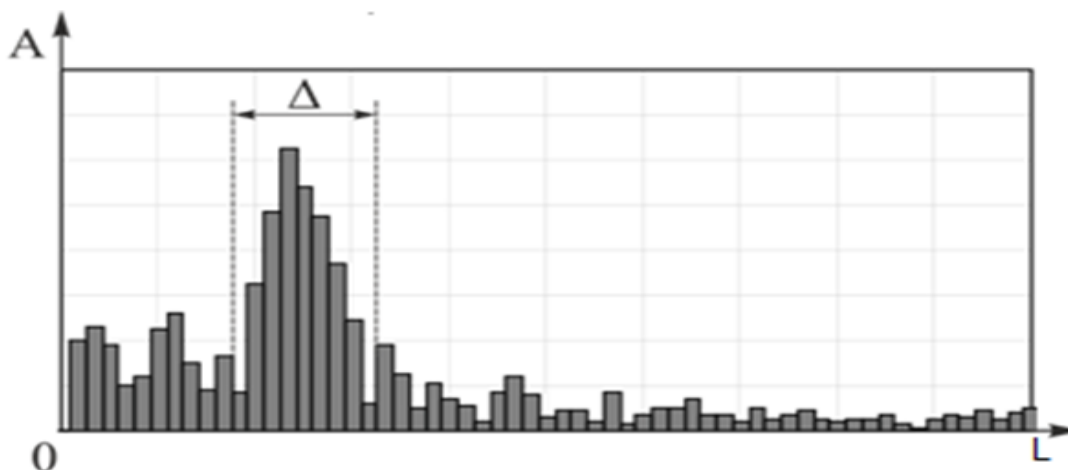


Figure 2. Amplitude spectrum of the frequency slice

To obtain a more accurate estimate of the heart rate, the harmonics adjacent to the maximum  $A_{jmax}$  are taken into account in the calculations, taking into account the decreasing nature of the near-maximum fragment of the amplitude spectrum before the first rises. This interval is indicated by the width  $\Delta$ . The contribution of these harmonics to the heart rate at the  $\Delta$ -range is taken into account by means of weighting coefficients  $w_\delta$  proportional to the amplitudes  $A_\delta$  of harmonics with the numbers  $\delta$  included in the  $\Delta$ -range of the amplitude spectrum of the signal of the frequency slice of the spectrogram, and  $\sum_{\delta \in \Delta} \omega_\delta = 1$ :

$$HR_f = \frac{60}{t_d \cdot n} \cdot \sum_{\delta \in \Delta} (w_\delta \cdot \delta),$$

where  $f \in [f_1, f_k]$  is the index of the frequency slice of the spectrogram of the original electrocardiogram,  $t_d$  is the sampling period of its samples,  $n$  is the length of the sliding window. The above formula for HRF is designed to measure heart rate in beats per minute.

Taking into account the weighting contribution of HRF, the final HR value is calculated by averaging HRF over all frequency indices  $f$  in the range  $[f_1, f_k]$ .

### Conclusion.

The proposed algorithm for heart rate measurement is implemented in a portable electrocardiograph based on an ESP32 microcontroller with Tensilica Xtensa LX6 architecture and an RC039 cardiomodule based on an AD8232 amplifier [12]. The average heart rate calculation time is no more than 5 ms for the length of the sliding window  $n = 128$  with a step of its movement equal to the length of the window and the size of the spectrogram matrix ( $n / 2 \times n / 8$ ). To control the heart rate, an analysis of the array of amplitude spectra obtained from the data of frequency slices of the spectrogram is used. The use of time-frequency conversion based on the window Fourier transform for processing this class of signals makes it possible to increase the accuracy of measuring heart rate during human movement.

### References

- [1] Islam, M. Development of Smart Healthcare Monitoring System in IoT Environment / Md. M. Islam, A. Rahaman, Md. R. Islam // SN Computer Science. – 2020. – № 1 (185). – P. 1 – 11. – [Электронный ресурс]. URL: <https://doi.org/10.1007/s42979-020-00195-y>.
- [2] Pradhan, B. IoT-Based Applications in Healthcare Devices / B. Pradhan, S. Bhattacharyya, K. Pal // Journal of Healthcare Engineering. – 2021. – Vol. 2021. – P. 1 – 18. - [Электронный ресурс]. URL: <https://downloads.hindawi.com/journals/jhe/2021/6632599.pdf>.

[3] Mahmoud, N. A Real-time Framework for Patient Monitoring Systems based on a Wireless Body Area Network / N. Mahmoud, S. El-Sappagh, S. M. Abdelrazek, H. M. El-Bakry // International Journal of Computer Applications. – 2020. – Vol. 176. – № 27. – P. 12–21. – [Электронный ресурс]. URL: <https://www.ijcaonline.org/archives/volume176/number27/mahmoud-2020-ijca-920274.pdf>.

[4] Sowmya, K. V. An efficeint health monitoring system with temperature and heart rate sensors using IoT / K. V. Sowmya, V. Teju // European Journal of Molecular & Clinical Medicine. – 2021. – Vol. 08. – № 2. – P. 793 – 802. – [Электронный ресурс]. URL: [https://ejmcm.com/article\\_7497\\_1eb61761da420cc16e28825b0ea9f268.pdf](https://ejmcm.com/article_7497_1eb61761da420cc16e28825b0ea9f268.pdf).

[5] Ghosh, A. Energy-Efficient IoT-Health Monitoring System using Approximate Computing / A. Ghosh, A. Raha, A. Mukherjee // Internet of Things. – 2020. – № 9. – P. 1–16. – [Электронный ресурс]. URL: <https://doi.org/10.1016/j.iot.2020.100166>.

[6] Осипов, А. Н. Архитектура системы дистанционного мониторинга состояния человека на основе технологии интернет вещей / А. Н. Осипов, А. П. Ключев, О. Ч. Ролич, С. А. Мигалевич, В. Д. Владымцев, И. О. Хазановский, А. В. Пацеев, С. В. Пацеев // Медэлектроника–2022. Средства медицинской электроники и новые медицинские технологии : сб. науч. ст. XIII Междунар. науч.-техн. конф. (Республика Беларусь, Минск, 8-9 декабря 2022 года). – Минск : БГУИР, 2022 – С. 40 – 42. – Режим доступа : [https://www.bsuir.by/m/12\\_100229\\_1\\_168906.pdf](https://www.bsuir.by/m/12_100229_1_168906.pdf).

## АЛГОРИТМ ИЗМЕРЕНИЯ ЧАСТОТЫ СЕРДЕЧНЫХ СОКРАЩЕНИЙ В СИСТЕМЕ ДИСТАНЦИОННОГО МОНИТОРИНГА СОСТОЯНИЯ ЧЕЛОВЕКА

**А. Н. Осипов**

Доцент кафедры электронной  
техники и технологии,  
Начальник Центра  
междисциплинарных  
исследований, к.т.н.

**О. Ч. Ролич**

Доцент кафедры  
проектирования  
информационно-компьютерных  
систем, к.т.н.

**М. П. Батура**

Профессор,  
заведующий НИЛ, д.т.н.

**А. П. Ключев**

Старший  
преподаватель  
кафедры инженерной  
психологии и  
эргономики

**В. Д. Владымцев**

Ассистент кафедры  
информатики

**С. А. Мигалевич**

Начальник ЦИИР,  
магистр технических  
наук

**Е. Р. Углонова**

Студент БГУИР

Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники», Республика Беларусь

E-mail: [osipov@bsuir.by](mailto:osipov@bsuir.by)

**Аннотация.** В статье рассмотрены вопросы измерения частоты сердечных сокращений при движении человека в системах удаленного мониторинга. Описан алгоритм измерения сердечных сокращений, основанный на обработке спектрограмм кардиограмм. Реализация алгоритма позволяет получить более точную количественную оценку частоты сердечных сокращений в реальном масштабе времени в условиях наличия помех и шумов.

**Ключевые слова:** частота сердечных сокращений, спектрограмма, электрокардиограмма, реальный масштаб времени.

UDC 004.93'1

## **DEEP LEARNING BASED APPROACH FOR VERTEBRAE DETECTION ON SPINE X-RAY**



**R.M. Charapennikau**  
*Student of the Department  
of Programming  
Technologies of the  
Belarusian State University  
fpm.cherepen@bsu.by*

**R.M. Charapennikau**

*Student of Faculty of Applied Math and Computer Science, BSU. Machine Learning Engineer at BGSoft.  
Research interests: medical imaging applications of deep learning, generative models.*

**Abstract.** In this paper two methods for vertebra detection on X-RAY images are presented. The key difference from standard object detection methods is that proposed ones (methods) are not limited to finding a rectangular bounding box, but can be used to detect an object with a bounding box of arbitrary quadrangle shape.

Our method achieves 0.988 Average Precision Score with 0.822 bounding box IoU.

**Keywords:** Deep learning, medical imaging, object detection, instance segmentation.

### **Introduction.**

Statistics indicate that there are around 540 million adults who experience back pain of different levels of intensity [1].

Traditional methods of diagnosing diseases of the spine are radiography, computed tomography (CT) and magnetic resonance imaging (MRI). Despite the fact X-Ray is a total projection image and is not so accurate compared to CT and MRI, it is still the most popular way of primary diagnosis. This is due to the following facts: X-Ray is fast (the whole procedure can be done in several minutes), cheap and no special patient preparation is required.

Accurately detecting and identifying each of the individual vertebrae is essential for medical professionals in diagnosing and treating spinal conditions such as spinal fractures, herniated discs, scoliosis and spondylolisthesis. Traditionally, spine vertebra detection was done manually by radiologists, which is a time-consuming and labor-intensive process.

At the same time Neural Networks have proven to be a reliable way to solve many problems related to image processing.

In this work we research the possibility of applying them to solve the problem of automatic detection of vertebrae from spine X-Ray.

On top of the detection methods, a diagnosis decision making support system for medical professionals can be built. Decision support system aims to help to make the diagnostic process faster, more accurate, and less labor-intensive for medical professionals compared to manual diagnostics.

### **Related Work.**

The object detection is the natural extension of object classification, which aims only at recognizing the object in the image. The goal of the object detection is to detect all instances of the predefined classes and provide its coarse localization in the image by axis-aligned boxes. The detector should be able to identify all instances of the object classes and draw a bounding box around it. It is generally seen as a supervised learning problem [2].

There are two major groups of object detection methods: Two-Stage Detectors (Faster-RCNN[3], Mask-RCNN[4]) and One-Stage Detectors (SSD[5], YOLO[6], RetinaNet[7]).

The key difference in these two types of object detection models is how they solve classification problem. Two stage models have separate region proposal subnetworks that generate bounding boxes, and later proposed regions are passed to classification subnetwork. At the same time one stage models solve region proposal and classification problems with a single subnetwork.

There also has been studies that explore automatic detection of spine diseases from X-Ray images. But they were limited to working in some particular regions of the spine, for example [8] is limited to Lumbar (S1, L1-L5 vertebrae, which is only 6 out of 32 vertebrae in the human body) spondylolisthesis.

In contrast, the goal of our research is to design a region-independent method of vertebrae detection, which can be further used to detect different pathologies.

#### **Model evaluation.**

Object detectors use multiple criteria to measure the performance of the detectors viz., frames per second (FPS), precision and recall. However, mean Average Precision (mAP) is the most common evaluation metric. Precision is derived from Intersection over Union (IoU), which is the ratio of the area of overlap and the area of union between the ground truth and the predicted bounding box. A threshold is set to determine if the detection is correct. If the IoU is more than the threshold, it is classified as True Positive while an IoU below it is classified as False Positive. If the model fails to detect an object present in the ground truth, it is termed as False Negative. Precision measures the percentage of correct predictions while the recall measures the correct predictions with respect to the ground truth.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

Based on the above equation, average precision is computed separately for each class. To compare performance between the detectors, the mean of average precision of all classes, called mean average precision (mAP) is used, which acts as a single metric for final evaluation [2].

#### **Vertebrae Detection.**

As described previously, in most cases, when detecting, it is sufficient to use as bounding boxes rectangles, whose sides are parallel to the coordinate axes. However, in the problem under consideration, they are not accurate enough. For example, in the task of determining scoliosis and its degree it is necessary that the sides of the bounding box be directed in the same way as the vertebra is rotated. At the same time the upper and lower sides are not necessarily parallel to each other. Therefore, we need to use an arbitrary convex quadrangle as a bounding box.

We propose two ways of achieving this:

- Finding a rectangular bounding with a standard object detection model followed by a refinement step using one more model, that solves regression tasks on 4 key points coordinates (method 1).
- Using an instance segmentation model to build a segmentation mask for each vertebra and future approximation of this mask with a quadrangle bounding box (method 2).

#### *Method 1*

At the first step of this approach, the problem of detecting individual vertebrae in the image using a rectangular bounding box. For this we can use any of the standard architectures for object detection task, for example, FasterR-CNN, YOLO, SSD.

In the second step, each of the obtained bounding boxes is cut out from the original image, and used as input data for the neural network that determines the coordinates of 4 key points of each vertebra.

Best results in detection of individual vertebrae on X-Ray images was achieved by FasterR-CNN architecture with ResNet-50 backbone.

Experiments were also carried out using one-step methods, such as SSD, RetinaNet (more

specifically, the experiments are described in Table 1). These methods, as expected, worked faster, but showed lower quality. Since in the problem under consideration the speed is not critical, preference was given to FasterR-CNN.

Table 1. Vertebrae detection experiments

<i>Architecture</i>	<i>Learning rate</i>	<i>Average Precision</i>	<i>Notes</i>
SSD (ResNet-50)	Starting: 3.0E-4 Decrease by 10 times every 5 epochs	Val: 0.921 Test: 0.918	ResNet-50 use pretrained ImageNet weights
RetinaNet(ResNet-50)	Starting: 3.0E-4 Decrease by 10 times every 5 epochs	Val: 0.948 Test: 0.949	ResNet-50 use pretrained ImageNet weights
Faster-RCNN (ResNet-50)	Starting: 3.0E-7 for backbone 3.0E-4 other layers Decrease by 10 times on plateau	<b>Val:</b> <b>0.9872</b> <b>Test:</b> <b>0.9884</b>	ResNet-50 use custom pretrained weights

One of the considered pathologies is scoliosis. To calculate the degree of scoliosis it is necessary to calculate the angles along the boundaries of the bounding boxes. Thus, one of the important characteristics of the bounding box is the coincidence of the angles between the lines of the correct and the predicted box. Therefore, it is proposed during training to optimize not only the absolute distance between points (as is usually done in regression problems), but also the angles between the lines. It is proposed to use  $AngleLoss = \sin^2(\theta)$  as a measure of the difference in angles, where  $\theta$  is angle between predicted and correct bounding box boundary. And as a generalization of this loss function to the bounding box, the average value over 2 pairs of lines (predicted and correct bottom and top boundaries) is used.

Table 2. Vertebrae keypoints regression experiments

<i>Architecture</i>	<i>Loss function</i>	<i>IoU</i>
ResNet-101	Mean squared error	Val:0.783 Test:0.794
ResNet-101	Mean absolute error	Val:0.766 Test:0.776
ResNet-101	Mean squared error + Angle Loss	<b>Val:0.811</b> <b>Test:0.822</b>
ResNet-101	Mean absolute error + Angle Loss	Val:0.786 Test:0.789

The results of applying method 1 to spine X-Rays are presented in Figure 1

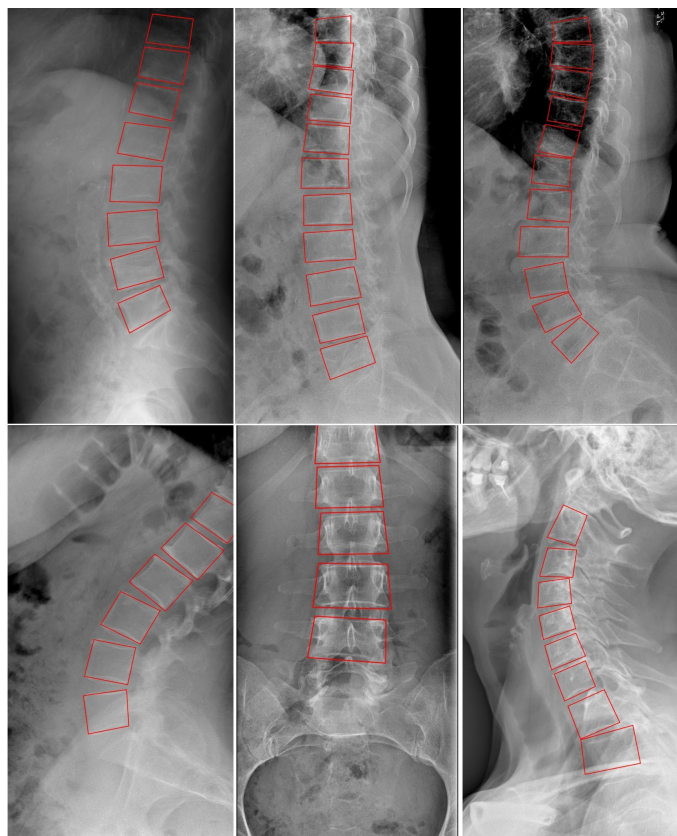


Figure 1. Results achieved with method 1

*Method 2*

The second approach for solving the problem under consideration is based on the instance segmentation. At the first step, masks are built for each of the vertebrae, and then they are approximated using quadrangles.

MaskR-CNN was used as an architecture for solving the instance segmentation problem. As a result, it was possible to achieve mask IoU **0.782** and bounding box Average Precision **0.984**.

To solve the problem of approximating masks using quadrangles, it is proposed to use gradient-free optimization methods.

The quadrangle is defined by 8 coordinates  $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ , using these coordinates we can build a mask that is an approximation of the resulting mask. And as an objective function, we use the IoU value between the mask obtained using MaskR-CNN and the mask built using the coordinates of the quadrangle. The following methods were chosen for comparison: the Nelder-Mead method, differential evolution, and the simulated annealing method (the results of the comparison of approximation methods are shown in Table 3).

Table 3. Mask approximation experiments

<i>Architecture</i>	<i>Mean <math>\pm</math> std execution time, s</i>	<i>Mean objective function value</i>
Nelder-Mead	0.0034 $\pm$ 0.008	0.925
<b>Differential evolution</b>	<b>0.1925 <math>\pm</math> 0.0386</b>	<b>0.965</b>
Simulated annealing	0.2047 $\pm$ 0.0074	0.960



The results of applying method 2 are presented on Figure 2.

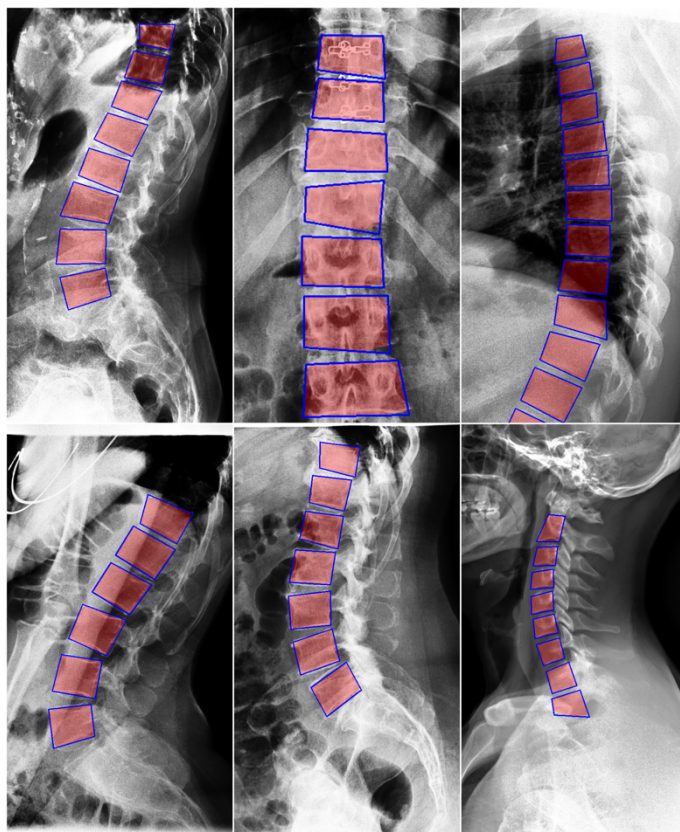


Figure 2. Results achieved with method 2

### Conclusion.

In this research we have explored the possibility of using neural networks for the automatic detection of individual vertebrae from spine X-Ray images.

We proposed two methods for achieving this, one using a standard object detection model followed by a refinement step using another model that solves regression tasks on 4 keypoints coordinates, and the other using an instance segmentation model to build a segmentation mask for each vertebra and future approximation of this mask with a quadrangle bounding box.

Our experiments showed that the FasterR-CNN architecture with ResNet-50 backbone produced the best results in detecting individual vertebrae on X-Ray images (**0.988** Average Precision with **0.822** bounding box IoU).

The proposed methods can help medical professionals in diagnosing and treating spinal conditions such as spinal fractures, herniated discs, scoliosis, and spondylolisthesis.

Further research can be done to improve the accuracy of the proposed methods and extend it to detect different pathologies.

### References

- [1] 57 Back Pain Statistics: How Common Is Back Pain? [Electronic resource]. URL: <https://www.crossrivertherapy.com/research/back-pain-statistics>
- [2] A Survey of Modern Deep Learning based Object Detection Models [Electronic resource]. URL: <https://arxiv.org/abs/2104.11892>
- [3] Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [Electronic resource]. URL: <https://arxiv.org/abs/1506.01497>
- [4] Mask R-CNN [Electronic resource]. URL: <https://arxiv.org/abs/1703.06870>

[5] SSD: Single Shot MultiBox Detector [Electronic resource]. URL: <https://arxiv.org/abs/1512.02325>

[6] You Only Look Once: Unified, Real-Time Object Detection [Electronic resource]. URL: <https://arxiv.org/abs/1506.02640>

[7] Focal Loss for Dense Object Detection [Electronic resource]. URL: <https://arxiv.org/abs/1708.02002>

[8] Detection of Lumbar Spondylolisthesis from X-ray Images Using Deep Learning Network [Electronic resource]. URL: <https://www.mdpi.com/2077-0383/11/18/5450>

## **ОСНОВАННЫЙ НА ГЛУБОКОМ ОБУЧЕНИИ ПОДХОД К ОБНАРУЖЕНИЮ ПОЗВОНКОВ НА РЕНТГЕНОГРАММЕ ПОЗВОНОЧНИКА**

***Р.М. Черепенников**  
Студент кафедры  
технологий  
программирования БГУ*

*Кафедры технологий программирования  
Факультет прикладной математики и информатики  
Белорусский государственный университет.  
Электронная почта: [fpm.cherepen@bsu.by](mailto:fpm.cherepen@bsu.by)*

**Аннотация.** В данной статье представлены два метода детекции позвонков на рентгеновских снимках. Основным отличием, предложенных методов, от существующих методов детекции является то, что они не ограничены поиском прямоугольной ограничивающей рамки, а решают задачу детекции с помощью ограничивающей рамки произвольной четырехугольной формы.

Предложенный метод достигает качества 0.988 метрики Average Precision при IoU ограничивающих рамок 0.822.

**Ключевые слова:** Глубокое обучение, методы медицинской визуализации, детекция объектов.



*Научное издание*

**BIG DATA И АНАЛИЗ ВЫСОКОГО УРОВНЯ**

**BIG DATA AND ADVANCED ANALYTICS**

СБОРНИК НАУЧНЫХ СТАТЕЙ  
IX МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ

В двух частях

Часть 2

(Республика Беларусь, Минск, 17–18 мая 2023 года)

В авторской редакции  
Ответственный за выпуск *Д. В. Лихачевский*  
Компьютерная верстка *И. В. Андриалович*

Подписано в печать 12.05.2023. Формат 60×84 1/8. Бумага офсетная. Гарнитура «Таймс».  
Цифровая печать. Усл. печ. л. 45,1. Уч.-изд. л. 42,5. Тираж 25 экз. Заказ 87.

Издатель и полиграфическое исполнение: учреждение образования  
«Белорусский государственный университет информатики и радиоэлектроники».  
Свидетельство о государственной регистрации издателя, изготовителя,  
распространителя печатных изданий №1/238 от 24.03.2014,  
№2/113 от 07.04.2014, №3/615 от 07.04.2014.  
Ул. П. Бровки, 6, 220013, г. Минск

