

BIG DATA

And Advanced Analytics

X Международная
научно-практическая
конференция

Часть 2



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
имени Мухаммада ал-Хоразмий, Узбекистан
Южный федеральный университет, Российская Федерация
ФГАОУ ВО «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
X МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ

(Республика Беларусь, Минск, 13 марта 2024 года)

В двух частях
Часть 2

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

Г.А. Пискун – кандидат технических наук, доцент, заместитель декана факультета компьютерного проектирования учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», Республика Беларусь

Рецензенты:

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

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

А. Косари – 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

BIG DATA и анализ высокого уровня = **BIG DATA and Advanced Analytics** : сб. науч. ст.
Б59 X Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 13 марта 2024 года). В 2 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. – Минск : БГУИР, 2024. – 495 с.
ISBN 978-985-543-752-0 (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 32.973.3я43

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



Джуманов Ж.Х., профессор Ташкентского университета информационных технологий имени Мухаммада ал-Хоразмий, доктор технических наук, профессор, Республика Узбекистан



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



Кругликов С.В., генеральный директор Государственного научного учреждения “Объединенный институт проблем информатики Национальной академии наук Беларуси



Лихачевский Д.В., декан факультета компьютерного проектирования БГУИР, кандидат технических наук, доцент, Республика Беларусь



Нестеренков С.Н., декан факультета компьютерных систем и сетей БГУИР, 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



Стемпцкий Р.А., проректор по учебной работе БГУИР, кандидат технических наук, доцент, Республика Беларусь



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



Тузиков А.В., заведующий лабораторией математической кибернетики Государственного научного учреждения «Объединенный институт проблем информатики Национальной академии наук Беларуси», доктор физико-математических наук, профессор, Республика Беларусь



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



Целых А.А., IT-директор – заместитель директора Института компьютерных технологий и информационной безопасности по информатизации Южного федерального университета (г. Таганрог), кандидат технических наук, доцент, Российская Федерация



Шнейдеров Е.Н., проректор БГУИР, кандидат технических наук, доцент, Республика Беларусь

ОРГАНИЗАТОРЫ КОНФЕРЕНЦИИ



Министерство образования Республики
Беларусь



Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»



Национальная академия наук Беларуси



Объединенный институт проблем
информатики



TASHKENT UNIVERSITY
OF INFORMATION
TECHNOLOGIES

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**

The logo for WhiteSnake, featuring the word "whitesnake" in a bold, lowercase sans-serif font. A red underline is positioned under the letter 'w'.

**Резидент Парка высоких технологий
WhiteSnake**

Быстро растущая IT-компания, специализирующаяся на Python разработке. Опытные разработчики создают изящный и чистый код для проектов в Fintech, FA&P, AI, Blockchain, EdTech, Medicine, Cybersecurity и Retail сферах. Клиентские проекты реализуются как для мировых брендов, так и для локальных заказчиков на рынках США, Европы и России.

ОГЛАВЛЕНИЕ

В.Ф. Алексеев, Е.И. Бавбель Оценка облаков точек БПЛА с помощью искусственного интеллекта	12
К.V. Tushynskaya, M.M. Zyranova, A.E. Asadchy Advancing BPM detection in hip-hop and R&B through audio representations and convolutional neural networks	19
А.А. Новицкий, Г.М. Шахлевич Лазерный комплекс для измерения параметров колёсных пар подвижного состава	26
Г.А. Пискун Моделирование сигналов в программной среде LabVIEW с возможностью экспорта и обработки данных в MS Excel	31
С.А. Мигалевич, А.Н. Марков, Г.А. Богрецов Внедрение интеграции обработки больших данных в строительстве	41
О.И. Лаврова Воздействие информационных и медиатехнологий на трансграничную мобильность капиталов	47
У.М. Козак, В.Г. Лунин Автоматизация процесса тестирования и отчетности на проекте Web-Insync в ЗАО “Альфа-Банк”	54
С.А. Мигалевич, А.Н. Марков, Д.Г. Ершов Интеграция обработки больших данных в автомобильной промышленности	66
А.Д. Хацкевич, В.Л. Ланин Моделирование локального индукционного нагрева для BGA и Flip-Chip монтажа	72
Г. Аннаныязова Обработка больших данных	78
Д. В. Дудко, М. В. Дашкевич Big Data в образовании: как использование данных помогает персонализировать учебный процесс и повышать успеваемость в школе	87
И.Б. Петухов, В.Л. Ланин Моделирование проволочного монтажа в многокристальных модулях	92
А.М. Тимофеев, А.Н. Шишпаренок, В.Е. Юреть Информационная безопасность больших данных на основе обезличивания персональных данных в соответствии с требованиями законодательства РБ	102
В.Н. Ильин, В.В. Баранов, О.В. Дударева Моделирование широкополосного ферромодуляционного магнитометра, индифферентного к хаотическому возбуждению чувствительного элемента	108
П.Р. Щербин Использование Big Data в производственной аналитике	117
С.М. Боровиков Статистическое имитационное моделирование в исследовании эффективности моделей прогнозирования надёжности изделий по информативным параметрам	122
И.А. Телеш, Е.А. Азаров, Д.С. Шумский, А.В. Горгун Система автоматического оповещения и контроля концентрации метана и углекислого газа в воздухе производственного помещения	132
М.В. Спринджук, В.И. Берник, В.С. Филинюк Актуальные вопросы преподавания технологий искусственного интеллекта, науки о данных и машинного обучения	138
О.Н. Шкор Использование генеративного искусственного интеллекта в автоматизации бизнес-процессов	141

Н.А. Кириенко Процедура выделения подсистем связанных функций для решения оптимизационных задач логического проектирования	145
Е.И. Ковшер, Д.Ю. Перцев Классификация и анализ снимков земной поверхности	152
И.П. Сидорчук, А.А. Охрименко, Е.Г. Крысь Электронное обучение для развития кадрового потенциала: опыт Российской Федерации	159
В.П. Корячко, А.В. Бакулев, М.А. Бакулева Алгоритм идентификации транспортных средств в условиях неполных данных	170
А.К. Куликов Изучение автоматизированных систем контроля и управления городскими фермами для разработки технической концепции системы управления городскими фермами в режиме онлайн	175
Н.А. Ларченко, И.В. Андриалович Big Data и Advance Analytics в персонализированной медицине	181
И.И. Ревинская, П.В. Камлач, Е.И. Лещевич Алгоритм обработки данных биомеханики дыхательных движений	187
Д.Н. Одинец, Е.А. Алуев Методика анализа больших массивов данных для оценки качества работы водителя и технического состояния автомобиля	194
М.Д. Михальков Преимущества и недостатки эмпирических моделей интеллектуальных систем поддержки принятия решений в области электронной коммерции	201
Е.И. Лещевич, П.В. Камлач, В.М. Бондарик, А.В. Чураков, Г.Д. Ситник, И.И. Ревинская Обработка больших данных при исследовании взаимодействия инфразвука с биологическими тканями	208
И.В. Марахина Бизнес-модели стартапов: методологические основы выбора	213
М.М. Лукашевич, В.Ю. Скобцов, А.В. Инютин, В.В. Ганченко Ансамбли нейронных сетей для классификации состояний и режимов функционирования подсистем космического аппарата	217
А.А. Сасновский, Р.А.Божко Big Data в автомобильной промышленности	229
Я.В. Лисовецкий, Т.Н. Дворникова Устройство управления исполнительными механизмами на микроконтроллере для умного дома	232
Т.М. Фильченкова Применение технологии смешанного обучения в учебном процессе учреждения высшего образования как способ повышения удовлетворенности студентов	238
В.А. Кабариха, О.Д. Семерник Нейронные сети для обнаружения лиц на изображениях	249
А.М. Gavrovska, А.В. Samčović, D.M. Dujković, Y.I. Golub, V.V. Starovoitov Weibull distribution based model behavior in color invariant space for blind image quality evaluation	254
А.А. Prihozhy, O.N. Karasik Blocked algorithm of shortest paths search in sparse graphs partitioned into unequally sized clusters	262
O.N. Karasik, А.А. Prihozhy Requirements to methods of graph clustering at the aim of solving the shortest path problem	272
П.А. Корбит, Я.Е. Шепелевский Этические и социальные аспекты использования Big Data в современном образовании	280
Е.И. Ковшер Метавселенная блокчейна	286
В.А. Реджепов, Д.Ю. Перцев Автоматическая диагностика сердечной аритмии во встроенных системах	290

М.В. Давыдов, Н.С. Давыдова, Д.Ю. Терещилко, Н.М. Елец Анализ динамических характеристик движений спортсменов с использованием динамометрической платформы	301
Д.И. Гусейнов, Н.С. Давыдова Кинематическая модель движений человека на базе матриц расстояний	308
А.Д. Каменкова Умные дома и Data Science: преобразование данных в производственных и бизнес средах. Преимущества и недостатки	315
Е.А. Лещенко, Е.В. Бегляк, В.А. Марцинкевич Методы и специфика защиты компьютерных сетей при работе с Big Data	320
A.N. Markov, M.M. Zyryanova, A.E. Asadchy Research of large language models for text generation and their practical application	325
D.A. Mirzaev Theoretical aspects of the development of banks in the conditions of digital transformation	335
Yang Shixin, A.A. Velchenko, S.A. Pauliukavets Motion control of mobile robot with differential drive and two castor wheels	340
Е.А. Алетурович, О.Д. Лосская, М.С. Ильясова, Ф.В. Усенко, Л.Р. Коркин, А.М. Прудник Инновационные технологии обучения для тренировки меткости стрельбы на примере VR-симулятора «Концентрация Мах»	346
М.В. Романюк, Е.А. Лещенко, А.Н. Марков Сетевая инфраструктура для работы с Big Data	354
К. Жаксылык, В. А. Захарьев Распределенная система потоковой обработки данных для задач распознавания речи	358
Ю.И. Голуб Оценка качества изображений после их сжатия алгоритмом jpeg	371
Ф.В. Усенко, М.С. Ильясова Оптимизация человеко-машинных взаимодействий через инновационные системы стимулирования трудовой деятельности	378
М.Н. Мельников, И.Д. Полегин, Ф.В. Усенко, М.С. Ильясова, Л.Р. Коркин, А.М. Прудник Приложение виртуальной реальности «Тур по музеям»	383
В.В. Старовойтов Следует ли нормализовать данные динамических подписей перед верификацией методом DTW?	391
Л.А. Демидова, Н.А. Морошкин Обучение вопросно-ответной нейросетевой модели на базе архитектуры модели LLaVA 1.5 с энкодером Saiga Mistral 7b и алгоритма низкоранговой адаптации LoRA	401
А.Ю.Саливон, Е.Ю.Красникевич, Г.О.Стешиц, Ф.В. Усенко, М.С. Ильясова, Л.Р. Коркин Игра виртуальной реальности «Островное исследование: головоломки и телепорты»	408
Д.Д. Лифарева, Е.А. Курьянова, Д.А. Веденькин Один из алгоритмов оценки стиля управления транспортным средством	413
Н.С. Сачивко, М.А. Калугина Алгоритмы повышения читабельности текста на изображениях низкого качества с применением нейросетевых моделей	418
Т.А. Васяева, Ю.А. Золушкин, А.В. Яковчук Разработка нейросетевого модуля прогнозирования страховых выплат	430
С.К. Корсакова, Н.В. Русина, В.М. Бондарик Прогнозирование временного ряда в Python с помощью метода Хольта-Уинтерса	438

Ван Там Лэ, С.М. Боровиков, С.К. Дик, А.В. Будник Ускоренная оценка надёжности разрабатываемых прикладных компьютерных программ.....	444
K.S Dzik, L.I. Piletskii, T.A. Asipovich Development of a method for using autoencoder to search for anomalies in cloud data.....	452
В.А Малицкий, А.Н. Осипов, А.П. Ключев, Тяньбо Ма Программный модуль обработки шумоподобных биомедицинских сигналов на основе коррелограммы.....	457
Н.К. Вахонин Моделирование и формирование пространственно-временного информационного обеспечения при информатизации, на примере отрасли мелиорация.....	460
В.Ф. Алексеев, Е.И. Бавбель Автономная посадка БПЛА с использованием визуального сервоуправления.....	472
С.А. Мигалевич, А.Н. Марков, А.Д. Стрельцов Обработка больших данных в туристической индустрии.....	480
М.В. Романюк, Е.А. Лещенко, С.С. Марковский Применение Big Data для защиты компьютерных сетей.....	486
Е.А. Бущик Исполняемое моделирование бизнес-процессов в системе Bizagi Process Modeler.....	490

УДК 004.021:004.75

ОЦЕНКА ОБЛАКОВ ТОЧЕК БПЛА С ПОМОЩЬЮ ИСКУССТВЕННОГО ИТЕЛЛЕКТА



В.Ф. Алексеев

Доцент кафедры проектирования
информационно-компьютерных систем
БГУИР, кандидат технических наук,
доцент
alexvikt.minsk@gmail.com



Е.И. Бавбель

Аспирант кафедры
проектирования информационно-
компьютерных систем БГУИР
egorigorevichw9@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.

Е.И. Бавбель

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации.

Аннотация. Разработки в области искусственного интеллекта достигли больших успехов в области автоматической семантической сегментации, как в 2D (изображении), так и в 3D пространствах. В контексте технологии 3D-записи она также нашла применение в нескольких областях, в первую очередь при создании облаков точек, что обычно выполняется вручную.

Сегодня изображения, полученные с помощью беспилотных летательных аппаратов (БПЛА), стали важным источником данных для исследователей, занимающихся картографированием в различных дисциплинах фотограмметрии и дистанционного зондирования. Реконструкция территории с помощью трехмерных (3D) облаков точек по изображениям, полученным с помощью БПЛА, является важным процессом, который можно использовать для традиционных двумерных кадастровых карт или для создания топографических карт. Облака точек следует классифицировать, поскольку они подвергаются различному анализу для извлечения дополнительной информации из прямых данных облака точек. Из-за высокой плотности облаков точек обработка данных и сбор информации делают классификацию облаков точек сложной задачей и могут занять много времени. Таким образом, обработка классификации позволяет найти оптимальное решение для получения ценной информации.

Ключевые слова: Облако точек, ИИ, БПЛА, ICP, FGR.

Введение. Благодаря передовым технологиям измерения, таким как наземные лазерные сканеры, требуется совсем немного усилий и времени для сбора миллионов и миллиардов 3D-точек в различных средах и для множества различных приложений. Однако большие коллекции 3D-данных сложно обрабатывать из-за отсутствия у них структуры. Это особенно актуально для

естественных сцен, где сложны не только сами захваченные данные, но и множество возможных вариаций объектов внутри сцены.

Большинство приложений избегают непосредственной работы с этими сложными наборами данных, передавая облака точек в различные представления, такие как треугольные сетки или более похожие на САПР структуры. Наиболее распространенной моделью САПР является конструктивная твердотельная геометрия, в которой параметрические трехмерные тела подгоняются к данным как примитивы формы. Такой подход быстро достигает своих пределов в сложных сценах, где большие части геометрии не покрываются библиотекой примитивов. Более общий подход заключается в извлечении контуров непосредственно из облаков точек и использовании их в качестве данных. На первый взгляд этот шаг может показаться малоэффективным, поскольку ребра также можно извлечь из треугольной сетки или параметрической модели. Чтобы избежать ошибок моделирования, которые часто возникают вокруг контуров, полезно знать контуры в качестве входных данных для триангуляции сетки или подгонки поверхности. Известно, что триангуляция сетки особенно сглаживает края [1].

Таким образом, интуитивно понятно, что контуры должны быть обнаружены до реконструкции поверхности, чтобы поддержать сегментацию. Системы интерактивного моделирования в графике работают именно таким образом. На практике обнаружить контуры в облаках точек сложно. Основными причинами являются не только неструктурированные данные, но и нечеткое определение того, что представляет собой контур. Хотя высокие значения кривизны играют важную роль, существуют и другие факторы, которыми нельзя пренебрегать. Например, кривизна на стыке двух стен здания в некоторых случаях может быть немного меньше, чем на неровных поверхностях, например, на лугах. Очевидно, что человек-оператор использует дополнительные контекстуальные подсказки, чтобы определить, что представляет собой контур. Эта ситуация, когда приходится использовать разрозненные знания, которые трудно сделать явными и формализованными, требует статистического машинного обучения.

Достижения в области технологий структурного движения (*SfM*) и беспилотных летательных аппаратов (БПЛА), широко известных как дроны, позволили создавать цифровые модели поверхности (*DSM*) с разрешением до миллиметра и сантиметра для количественной оценки поверхностных процессов, которая будет использоваться в качестве базовой карты для геологического картирования, определения структуры растительности, исследования эрозии почвы, изучения городских особенностей. *SfM* разработан в области изучения компьютерного зрения для трехмерной (*3D*) реконструкции объектов на основе сильно перекрывающихся фотографий, сделанных с помощью широко используемых камер, смартфонов и планшетов. *SfM* значительно снижает стоимость топографической съемки, поскольку традиционная фотограмметрия требует специальных и дорогих камер и самолетов для реконструкции как природных, так и искусственных особенностей местности. В частности, интеграция *SfM* с недорогими БПЛА дает возможность проводить крупномасштабные реконструкции поверхности для географических и экологических исследований.

Обычная трехмерная топографическая реконструкция на основе *SfM*-БПЛА включает в себя два этапа. Первый этап – сделать серию фотографий камерами БПЛА. Эти фотографии затем используются для создания облака точек, цифровой модели поверхности нанесенной на карту территории и ортофотоплана, чтобы обеспечить детальное и точное представление топографии местности.

Регистрация и точность крупномасштабных географических данных. Процесс регистрации математически представляет собой преобразование координат на основе матрицы преобразования, представляющей изменения смещения, вращения и масштаба. Пусть P – облако точек, которое необходимо выровнять, а Q – облако опорных точек. Значит P_c и Q_c – облако точек, созданные из опорных объектов, выбранных из P и Q соответственно. К выбору эталонного объекта предъявляется несколько требований:

1. Опорный объект должен быть зафиксирован в нужном положении;

2. Область, ограниченная эталонным объектом, должна включать целевой объект для регистрации.

3. Внешний вид и расположение эталонного объекта не изменятся со временем.

Предположим, регистрация выполнена на P_C чтобы выровнять его по Q_C и результирующая матрица преобразования задается как T . Тогда мы имеем следующее соотношение:

$$T(P_C)=Q_C$$

Предполагая, что облако точек сформировано правильно, область, ограниченная P_C такое же, как и заключенное в Q_C . Поскольку целевые объекты из обоих наборов данных облака точек одинаковы, местоположение P_T в P_C совпадает с расположением Q_T в Q_C . Более того, поскольку T линейно, имеем:

$$T(P_T)=T(Q_T)$$

Матрица преобразования, созданная путем регистрации опорных объектов, может правильно выровнять целевой объект. Следовательно, вместо того, чтобы выполнять регистрацию всего облака точек, мы можем выполнить регистрацию только выбранных стабильных объектов, чтобы получить матрицу преобразования, а затем применить ее ко всему облаку точек.

Как показано на рисунке 1, весь процесс был разделен на две части. Первая часть заключается в создании 3D-облаков точек и 2D-ортофотоснимков различных съемок на основе программного обеспечения *SfM*, такого как *Pix4D* и *Photoscan*. Вторая часть – это предлагаемый рабочий процесс для совместной регистрации облаков точек различных съемок. Эта часть состоит из трёх компонентов [3].

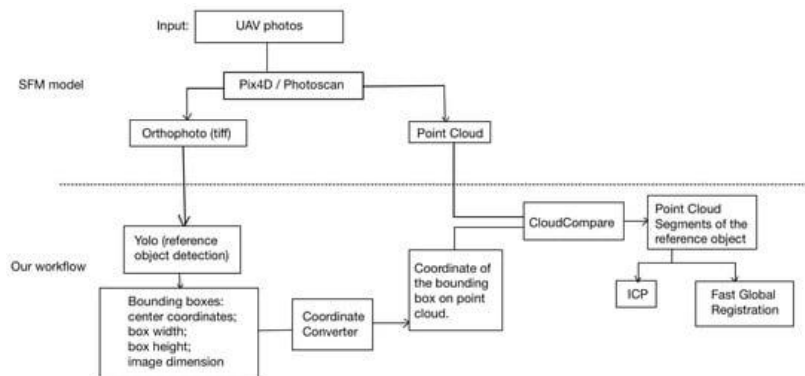


Рисунок 1. Полный рабочий процесс предлагаемой модели.

Первый заключается в извлечении стабильных эталонных объектов из ортофотоснимков, полученных с помощью фотографий, собранных БПЛА. Именуемой моделью обнаружения эталонных объектов [3].

Второй компонент определяет координаты ограничивающей рамки опорных объектов в облаке точек. Эти координаты используются для вырезания облаков точек опорных объектов из всего облака точек – моделью преобразователя координат.

Последний компонент предназначен для регистрации обрезанных облаков точек опорных объектов, а полученная матрица преобразования затем применяется ко всему облаку точек – моделью регистрации сегментов [4].

Модель регистрации сегментов. При регистрации в облаке в основном применяются две существующие модели: *ICP* и *FGR*. Оба алгоритма удобно интегрируются в *CloudCompare*. Однако крайне важно понимать принципы, лежащие в основе этих методов, и эффективно корректировать данные и параметры, чтобы улучшить время работы и точность регистрации.

Существуют разработанные комплексные рабочие процессы для алгоритмов *FGR* (рисунок 4) и *ICP* (рисунок 5), стремясь повысить ясность и понятность каждого этапа процесса регистрации. Что касается размера данных, эти два метода работают по-разному в зависимости от своих принципов.

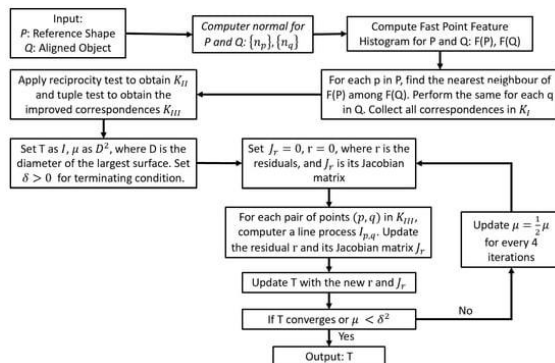


Рисунок 4. Рабочий процесс *FGR*.

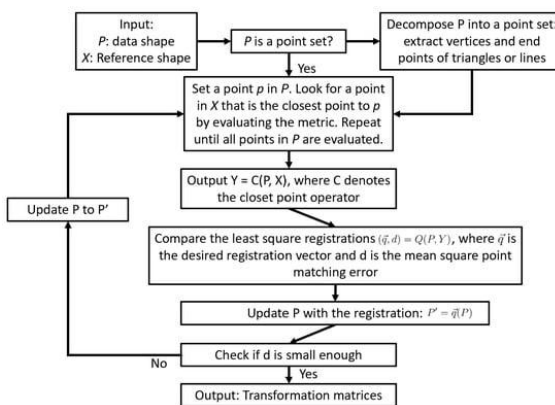


Рисунок 5. Рабочий процесс итеративных ближайших точек.

Метод *FGR* известен своей исключительной вычислительной эффективностью, особенно при работе с большими облаками точек. В отличие от других алгоритмов регистрации с потенциально более высокой временной сложностью, метод *FGR* поддерживает линейную временную сложность в порядке точек ($O(n)$). Это преимущество делает его особенно подходящим для регистрации крупномасштабных облаков точек. Более того, итеративность и параллельность данных метода *FGR* делают его легко поддающимся распараллеливанию, используя возможности многоядерных процессоров или ускорение графического процессора. Эта возможность параллельной обработки значительно ускоряет процесс регистрации, делая метод *FGR* еще более подходящим для крупномасштабных задач выравнивания облаков точек. В результате метод *FGR* является идеальным выбором при работе с обширными наборами данных, предлагая как вычислительную эффективность, так и преимущества распараллеливания.

ICP, напротив, работает лучше, когда нужно сопоставить и оптимизировать меньше точек. *ICP* быстрее сходится к приемлемому выравниванию. На каждой итерации алгоритм стремится минимизировать расстояние между соответствующими точками, а при меньшем количестве точек этот процесс становится более эффективным, что приводит к более быстрой сходимости. Временная сложность *ICP* обычно линейна относительно количества точек в наборах данных. Следовательно, когда существует меньше точек, вычислительная нагрузка алгоритма снижается, что делает его более эффективным для небольших облаков точек. Кроме того, поиск соответствий между двумя облаками точек является важным шагом в алгоритме *ICP*. При меньшем количестве

точек поиск соответствующих точек становится менее сложным и трудоемким, что повышает точность выравнивания. Качество первоначальной оценки выравнивания также может повлиять на результат *ICP*.

Основываясь на вышеизложенном, модель регистрации сегмента выполняет функцию определения соответствующего метода регистрации, который следует применить, путем оценки атрибутов обрезанного облака точек. Рекомендуется, чтобы, когда обрезанное облако точек содержит относительно небольшое количество точек, лучшим выбором для регистрации был *ICP*. И наоборот, если обрезанное облако точек имеет значительный размер, лучшим вариантом будет *FGR*. Хотя временная сложность методов *ICP* и *FGR* остается $O(n^2)$ и $O(n)$ соответственно, размер выборки n значительно уменьшается за счет простого использования облаков точек эталонных объектов. Следовательно, время работы может быть улучшено.

Испытание методов *ICP* и *FGR*. Согласно результатам испытаний, рассмотренным в [5], наиболее эффективным является применение метода *ICP* на частичном облаке точек черно-белых мишеней. Он превосходит другие как по эффективности использования времени, так и по точности регистрации, при этом среднее расстояние меньше, чем среднее расстояние между точками в наборах данных. Эти регистрационные мишени размещались в одних и тех же точках крепления во время каждого исследования, что является относительно стабильными данными. Дополнительная очистка облаков точек может повысить точность регистрации. Кроме того, использование только черно-белых целей уменьшило размер облаков точек для регистрации, что снизило вероятность возникновения проблемы с локальными минимумами. Напротив, ошибка, возникающая при применении *ICP* ко всему облаку точек, примерно в 600 раз превышает ошибку регистрации, применяемой исключительно к черно-белым объектам. Это указывает на то, что применение метода *ICP* к большому набору данных облака точек не является стабильным.

Что касается метода *FGR*, то он работает значительно быстрее в рамках предложенного рабочего процесса с разумной ошибкой. Это преимущество становится особенно важным при работе с большими наборами данных, где время обработки играет решающую роль в реальных приложениях. Применение метода *FGR* с предложенным рабочим процессом приводит к меньшей ошибке по сравнению с регистрацией с использованием метода *ICP* для всего облака точек. Возможно, что применение метода *FGR* ко всему облаку точек может дать несколько более высокую точность. Однако обработка регистрации всего набора данных может занять значительно больше времени. Таким образом более практично применить метод *FGR* в рабочем процессе [5].

В тестовых случаях черно-белые мишени оказались оптимальным выбором для эталонных объектов, что позволяет успешно применять метод *ICP* в рабочем процессе. Это связано с их фиксированным позиционным характером и пригодностью к сегментации, снижающей негативное влияние шумов и нецелевых точек на регистрацию. Однако важно отметить, что этот подход зависит от наличия таких объектов. В сценариях, где такие объекты отсутствуют, метод *FGR* может быть более надежным при регистрации [5].

Данные испытаний были собраны на относительно плоском склоне холма. Однако можно считать, что этот рабочий процесс на основе искусственного интеллекта может применяться и к другим топографическим условиям, например, к крутому рельефу, при условии, что можно обнаружить устойчивые наземные целевые объекты. Матрица преобразования генерируется на основе регистрации частичных облаков точек обнаруженных стабильных объектов, на которые не влияет крутой рельеф. Можно использовать простой подход с Z -порогом для удаления шума из-за его эффективности в наших тестовых данных. Этот метод удаления шума может не подойти для случаев, связанных с крутой топографией, поскольку вариация значений z становится намного больше. Этот этап удаления шума является необязательным, и пользователи могут реализовать другие методы удаления шума в этих топографических настройках [5].

Несмотря на хорошую производительность модели при регистрации больших наборов данных, для расширения ее возможностей целесообразно выполнить некоторые улучшения. В-первых, модель обнаружения объектов можно дополнительно обучить, чтобы она могла

обнаруживать более широкий спектр объектов, которые могут существовать в различных топографических условиях. В настоящее время модель обучена обнаруживать только три типа объектов, а именно: черно-белую цель для БПЛА, трубы и дома (палатки). Однако эти цели могут не появиться в других областях регистрации. Кроме того, можно использовать методы параллельных вычислений для оптимизации эффективности нашей модели. Благодаря мощности параллельной обработки мы можем значительно ускорить выполнение таких задач, как удаление шума, одновременно снижая риск переполнения буфера. Можно сделать этот рабочий процесс более доступным и удобным для пользователя, создав в *CloudCompare* плагин с более настраиваемыми интерфейсами. В настоящее время пользователям приходится вручную применять результирующую матрицу преобразования для регистрации всего облака точек, что может занять много времени. Автоматизация всех шагов, связанных с этим рабочим процессом на основе искусственного интеллекта, может упростить процесс регистрации и значительно повысить удобство и эффективность этого рабочего процесса.

Заключение. Алгоритмы итерационной ближайшей точки (*ICP*) и быстрой глобальной регистрации (*FGR*) являются распространенными подходами, используемыми для регистрации облака точек. Однако они не предназначены специально для обработки больших наборов данных облаков точек. Чтобы устранить это ограничение, необходимо разработать рабочий процесс на основе искусственного интеллекта для извлечения стабильных опорных объектов, а затем использовать только частичные облака точек из стабильных опорных объектов, чтобы обеспечить точную и эффективную регистрацию данных большого облака точек. Этот рабочий процесс обеспечивает более контролируемый и целенаправленный процесс регистрации, специально адаптированный к интересующим эталонным объектам. Такой целенаправленный подход сводит к минимуму возможность ошибок или неточностей, способствуя более надежной и надежной регистрации.

Одним из ключевых преимуществ предлагаемого в данной статье рабочего процесса является его гибкость. Пользователям разрешено выбирать собственные эталонные объекты для процесса регистрации. Включив методы *ICP* и *FGR* в предлагаемый нами рабочий процесс и применив их к меньшим сегментам облака точек, мы можем эффективно выравнивать облака точек, одновременно снижая риск возникновения проблем с локальными минимумами.

Список литературы

- [1] July 2012 ISPRS Annals of the Photogrammetry Remote Sensing and Spatial Information Sciences I-3 DOI:10.5194/isprsannals-I-3-245-2012
- [2] Eltner, A.; Schneider, D. Analysis of different methods for 3D reconstruction of natural surfaces from parallel-axes UAV images. *Photogramm. Rec.* 2015, 30, 279–299. [Google Scholar] [CrossRef].
- [3] Fitzgibbon, A.W. Robust registration of 2D and 3D point sets. *Image Vis. Comput.* 2002, 21, pp. 1145–1153.
- [4] Jost, T.; Hügli, H. Fast ICP algorithms for shape registration. In *Proceedings of the 24th DAGM Symposium, Zurich, Switzerland, 16–18 September 2002*; pp. 91–99.
- [5] Feng Y, Leung KL, Li Y, Wong KL. An AI-Based Workflow for Fast Registration of UAV-Produced 3D Point Clouds. *Remote Sensing.* 2023; 15(21):5163. <https://doi.org/10.3390/rs15215163>

Авторский вклад

Алексеев Виктор Федорович – руководство исследованием принципов, лежащие в основе методов *ICP* и *FGR*.

Бавбель Егор Игоревич – постановка задачи исследования, описание принципа работы алгоритмов *ICP* и *FGR*, их сравнения и изучения применимость с БПЛА системами.

ASSESSMENT OF UAV POINT CLOUDS USING ARTIFICIAL INTELLIGENCE

V.F. Alexseev

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

E.I. Bavbel

*Postgraduate student, Department
of Information Computer Systems
Design BSUIR*

Abstract. Developments in the field of artificial intelligence have made great strides in the field of automatic semantic segmentation, both in 2D (image) and 3D spaces. In the context of 3D recording technology, it has also found application in several areas, most notably in the creation of point clouds, which is usually done manually.

Today, unmanned aerial vehicle (UAV) imagery has become an important source of data for mapping researchers in various disciplines of photogrammetry and remote sensing. Reconstructing an area using three-dimensional (3D) point clouds from UAV imagery is an important process that can be used for traditional 2D cadastral maps or for creating topographic maps. Point clouds should be classified as they undergo various analyzes to extract additional information from the direct point cloud data. Due to the high density of point clouds, data processing and information collection make point cloud classification challenging and can be time consuming. Thus, classification processing allows finding the optimal solution to obtain valuable information.

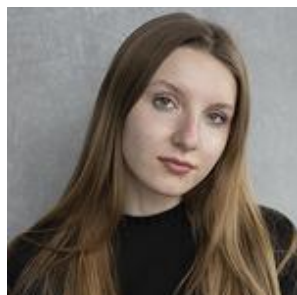
Keywords: Point cloud, AI, UAV, ICP, FGR.

UDC 004.021:004.75

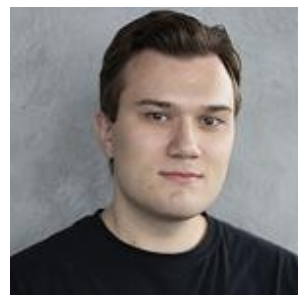
ADVANCING BPM DETECTION IN HIP-HOP AND R&B THROUGH AUDIO REPRESENTATIONS AND CONVOLUTIONAL NEURAL NETWORKS



K.V. Tushynskaya
Teacher assistant of the Faculty
of Computer Systems and
Networks BSUIR, master degree
student
katetushkan@icloud.com



M.M. Zyranova
Last year student of the
Faculty of Computer
Systems and Networks
BSUIR
z.y.r.y.a.n.o.v.a@mail.ru



A.E. Asadchy
Last year student of the Faculty
of Computer Systems and
Networks BSUIR
aleh.asadchy@gmail.com

K. Tushynskaya

Graduated from the Belarusian State University of Informatics and Radioelectronics. The areas of scientific interests are related to the investigation and analyzing of different music representations and their ability to give accurate information about the track.

M. Zyranova

Studying at the Belarusian State University of Informatics and Radioelectronics at the fourth course. The areas of scientific interests are related to the development and implementation of convolutional neural networks for music beats detection and analyzing.

A. Asadchy

Studying at the Belarusian State University of Informatics and Radioelectronics at the fourth course. The areas of scientific interests are related to the development and implementation of convolutional neural networks for music beats detection and analyzing.

Abstract. This study introduces a method for improving BPM detection in hip-hop and R&B music by integrating audio representations with Convolutional Neural Networks (CNNs). Through analyzing tempograms, spectrograms, and onset features optimized for CNN processing, our approach demonstrates enhanced accuracy in BPM detection across these genres. Evaluated on the Million Song Dataset (MSD), our findings offer significant advancements for automated music analysis and applications in music recommendation and genre classification.

Keywords: Beats per Minute (BPM), Convolutional neural network (CNN), tempogram, beatgram, rhythmogram, tempo-invariant processing.

Introduction. The dynamic and complex nature of contemporary music, particularly within the hip-hop and R&B genres, presents unique challenges in the field of music information retrieval (MIR). These genres are renowned for their intricate rhythms, diverse tempos, and the prominent use of vocal elements, all of which play a pivotal role in defining their sonic identity. Accurately detecting Beats Per Minute (BPM) is crucial for analyzing, categorizing, and interacting with these musical forms, yet the variability and complexity inherent in hip-hop and R&B tracks often elude traditional BPM detection methods.

This paper introduces a focused exploration into different audio representations as a solution to improve BPM detection specifically within hip-hop and R&B music. By examining a variety of audio representations, including tempograms, rhythmograms, chromagrams, spectrograms, and onset and tempo features, we aim to identify those that are most effective in

capturing the unique rhythmic and melodic characteristics of these genres. The selected representations are further optimized for integration with Convolutional Neural Networks (CNNs), enhancing their ability to accurately detect BPM across a wide range of hip-hop and R&B compositions.

Our investigation not only addresses the technical challenges associated with BPM detection in these genres but also seeks to contribute to the broader understanding of their musical structure through advanced audio analysis techniques. By narrowing our focus to hip-hop and R&B, we aim to develop methodologies that are finely tuned to the nuances of these genres, thereby offering more precise and reliable tools for music analysis, genre classification, and the creation of engaging music.

Audio Representation Comparison Analysis. In our dedicated pursuit to refine and enhance the accuracy of Beats Per Minute (BPM) detection methodologies, particularly within the nuanced soundscapes of hip-hop and R&B music, we embarked on a comprehensive reevaluation of the spectrum of audio representations.

Our research endeavors have led us to place a significant emphasis on the examination and integration of rhythmograms and chromagrams, alongside the established utility of tempograms, spectrograms, and onset and tempo features. This strategic recalibration in our focus is driven by the recognition of the paramount importance of vocal and melodic elements in these genres, which often serve as the primary indicators of BPM.

Hip-hop and R&B, genres celebrated for their rich lyrical density and melodic complexity, demand a nuanced approach to BPM detection that goes beyond the conventional metrics. The rhythmic cadence of spoken word, the syncopation of beats, and the harmonic undercurrents present unique challenges and opportunities for audio analysis.

By including rhythmograms and chromagrams in our analytical arsenal, we aim to capture the essence of these elements more effectively. This initiative not only reflects our commitment to addressing the specific needs of hip-hop and R&B music analysis but also underscores our broader objective to advance the precision of BPM detection technologies.

Through this refined approach, we seek to contribute meaningful insights and methodologies that can adapt to the evolving landscapes of music genres, thereby enhancing the applicability and accuracy of BPM detection in music information retrieval systems.

Below is the comparison table 1 that summarizes our findings from the comparative analysis of the selected audio representations. This table highlights the pros and cons of each representation and indicates their selection for further investigation based on their potential to enhance BPM detection in hip-hop and R&B music.

The criteria for evaluation include the representation's ability to capture genre-specific rhythmic and melodic nuances, computational efficiency, and adaptability to convolutional neural network architectures for automated BPM detection.

Table 1. Comparative Analysis of Audio Representations

Representation	Function	Strengths	Weaknesses
Tempogram	Detecting and analyzing rhythmic patterns in music.	Emphasis on rhythmic periodicities	Sensitive to changes in timbre and instrument characteristics.
Rhythmogram	Detecting and analyzing beats and rhythm-related features in music.	Stable representations for detecting beats and rhythm-related features.	Can be computationally expensive.
Spectrogram	Analyzing the frequency content and temporal evolution of sounds	Unique information about its frequency content.	Temporal evolution tools like spectral filtering and morphing can be used for creative sound design
Chromagram	Identifying harmonic patterns and melodic characteristics	Similar to a spectrogram, but only shows the intensity of the twelve pitches that make up the equal-tempered scale.	Often used in music analysis
Onset and Tempo Features	Creating representations that are less sensitive to tempo variations.	Less sensitive to tempo variations when processed properly	May miss subtle rhythmic nuances.
Percussive and Harmonic Separation	Making the representation more robust to tempo changes.	Isolates rhythmic elements	Can be difficult to achieve perfect separation.
Pulse or Beat Tracking	Providing a stable foundation for analyzing rhythmic patterns.	Provides a stable foundation for analyzing rhythmic patterns	Can be inaccurate in complex or polyrhythmic music

This table presents the strengths and limitations of each audio representation, with all selected for further investigation to leverage their unique contributions towards enhancing BPM detection in hip-hop and R&B genres

Given the unique rhythmic and melodic characteristics of hip-hop and R&B, rhythmograms and chromagrams have been reintegrated into our analysis framework alongside tempograms, spectrograms. This selection aims to harness the strengths of each representation to accurately detect BPM, reflecting the complex interplay of beats, melodies, and vocals inherent in these genres.

Exploiting CNNs for Hierarchical Rhythm Analysis in Music. Convolutional Neural Networks (CNNs) have revolutionized the field of machine learning, offering robust frameworks for extracting and learning from hierarchical data structures. Originally popularized within the realm of image analysis, the principle behind CNNs – that of learning layered feature representations – translates effectively to the domain of audio analysis, particularly for the study of rhythm in music.

In rhythm analysis, CNNs exploit their hierarchical structure to dissect and understand the complex layers of musical composition. The initial layers of a CNN are adept at identifying fundamental rhythmic elements, such as individual beats or onset events, which can be thought of as the auditory equivalents of edges in image analysis. These basic components are the building blocks of rhythm, providing the tempo and the groove of the track.

As the information progresses through the network, intermediate layers amalgamate these basic components into more sophisticated rhythmic patterns. In the context of music, this might include the identification of rhythmic motifs, syncopation, or the interplay between different rhythmic instruments. These patterns are akin to the shapes and textures that CNNs learn to recognize in images, offering a deeper understanding of the music's rhythmic structure.

The efficacy of CNNs in music rhythm analysis is significantly enhanced by incorporating multi-dimensional audio representations. Spectrograms, for instance, offer a rich, time-frequency depiction of music, allowing CNNs to capture not only rhythm but also pitch and timbre information. Meanwhile, tempograms and rhythmograms provide targeted insights into temporal dynamics and rhythmic consistency, crucial for genres that thrive on rhythmic innovation.

This integration enables CNNs to not only recognize but also predict rhythmic patterns, facilitating advancements in automated music generation, remixing, and interactive music systems. It underscores the transformative potential of CNNs in music technology, bridging the gap between computational analysis and creative musical expression.

While the primary focus here is on rhythm analysis, the capability of CNNs extends to capturing harmonic and melodic features when trained on appropriate representations like chromagrams. This holistic approach to music analysis underscores the versatility of CNNs, enabling a comprehensive understanding of music beyond the rhythmic dimension.

CNN architecture for rhythm analysis. The architecture of our Convolutional Neural Network (CNN) for rhythm analysis is a multi-layered structure carefully designed to uncover intricate rhythmic patterns in music. The journey through this network begins with the input layer, represented as 2D matrices derived from tempograms, rhythmograms, or beatgrams. These representations encapsulate the temporal dynamics of the music and serve as the foundation for further analysis.

The initial convolutional layer embarks on the task of capturing local rhythmic features. It employs a sequence of convolutional layers, each equipped with filters of varying sizes to discern different rhythmic patterns. At this stage, batch normalization and activation functions (such as Rectified Linear Units – ReLU) are applied after each convolution, ensuring that the learned features are fine-tuned to highlight relevant rhythmic elements. Subsequently, max-pooling layers come into play, strategically downsampling the learned features while preserving essential rhythmic information.

As the data flows deeper into the network, additional convolutional layers are introduced to extract abstract and higher-level rhythmic patterns. The number of filters in these layers progressively increases to accommodate the growing complexity of the learned features. This hierarchical approach allows the network to discern intricate rhythmic nuances that may be deeply embedded within the music.

To summarize the hierarchical features learned across the entire representation, global average pooling is applied. This step effectively condenses the wealth of rhythmic information into a manageable form, facilitating subsequent analysis. The network then connects the global average pooling layer to a sequence of fully connected layers. To prevent overfitting and ensure robustness, the number of neurons in each layer is gradually reduced.

The final output layer is tailored to the specific task at hand, which may include BPM range prediction or genre classification. For classification tasks, a softmax layer is employed to predict a certain class label, offering insights into the rhythmic characteristics or genre affiliation of the analyzed music.

This CNN architecture is meticulously crafted to decode the intricate rhythmic fabric of music, gradually progressing from local features to more abstract patterns, and finally providing valuable insights through classification. It exemplifies the power of deep learning in unraveling the complexities of rhythm analysis.

Task results. For our research, we selected the Million Songs Dataset (MSD) as the primary source of information. The MSD is a comprehensive and freely-available collection of audio features and metadata for a vast repository of contemporary popular music tracks. At its core, the dataset encompasses feature analysis and metadata for an astounding one million songs, making it an invaluable resource for music-related research.

To extract the audio data from the MSD, we utilized the Python library librosa. Librosa is a versatile library tailored for audio and music processing, enabling us to access and manipulate the audio data effectively. Through librosa, we were able to represent each track as a 2D array of numerical values, capturing essential audio features for our analysis.

As part of our investigation into improving BPM detection in hip-hop and R&B music, we selected four distinct representation types for analysis. These representations serve as the foundational elements for our neural network's training dataset. The examples of representations are displayed at figure 1.

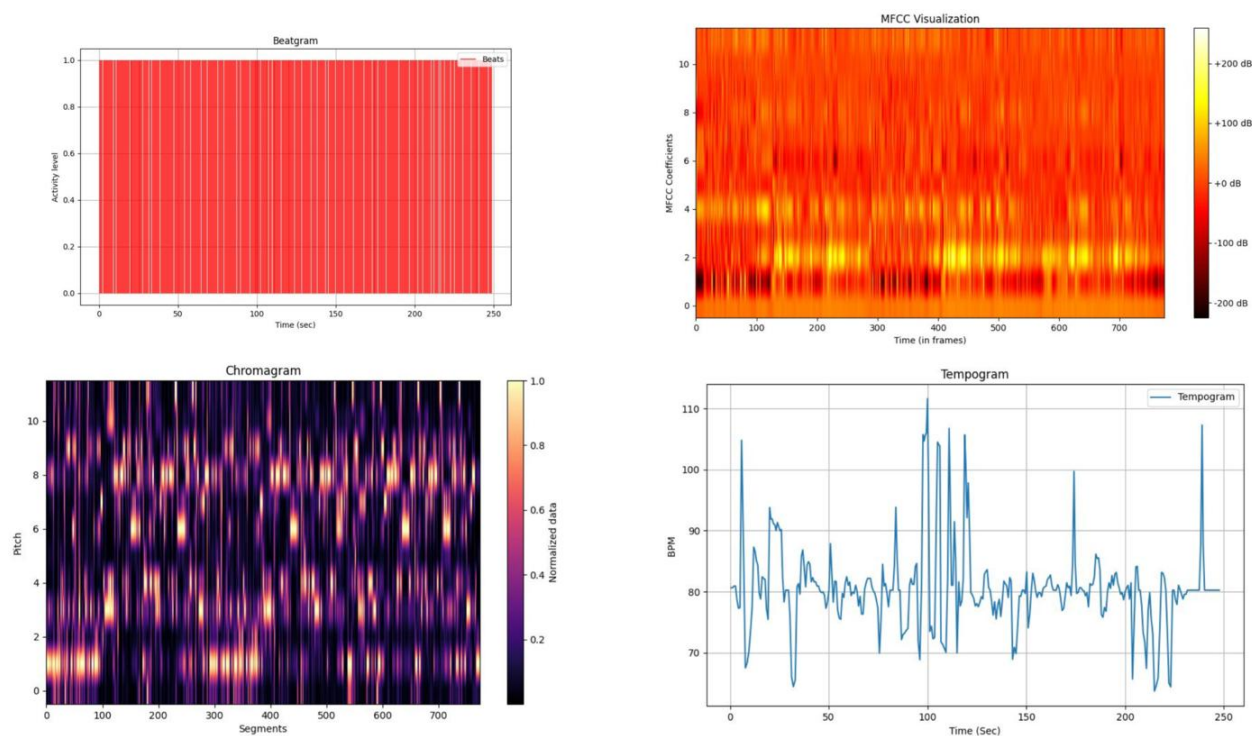


Figure 1. Examples of audio representation in 2d arrays: beatgram, chromagram, spectrogram, tempogram

In our endeavor, we assembled a dataset consisting of 10,000 plots for each of the four selected representation types: tempograms, beatgrams, chromagrams, and spectrograms. This dataset was thoughtfully divided into both training and testing subsets to facilitate the training and evaluation of our Convolutional Neural Network (CNN).

Following the dataset preparation, our CNN underwent extensive training over 1,000 epochs. During this training process, we closely monitored the learning curves to discern the representation type that yielded the most accurate results. It's important to note that the primary goal of this investigation was to establish correlations between different audio representations.

The outcome of this rigorous analysis is presented below at figure 2, shedding light on the performance of various representation types and their impact on the accuracy of our neural network. The plot unequivocally demonstrates that the tempogram representation outperforms others in terms of accuracy.

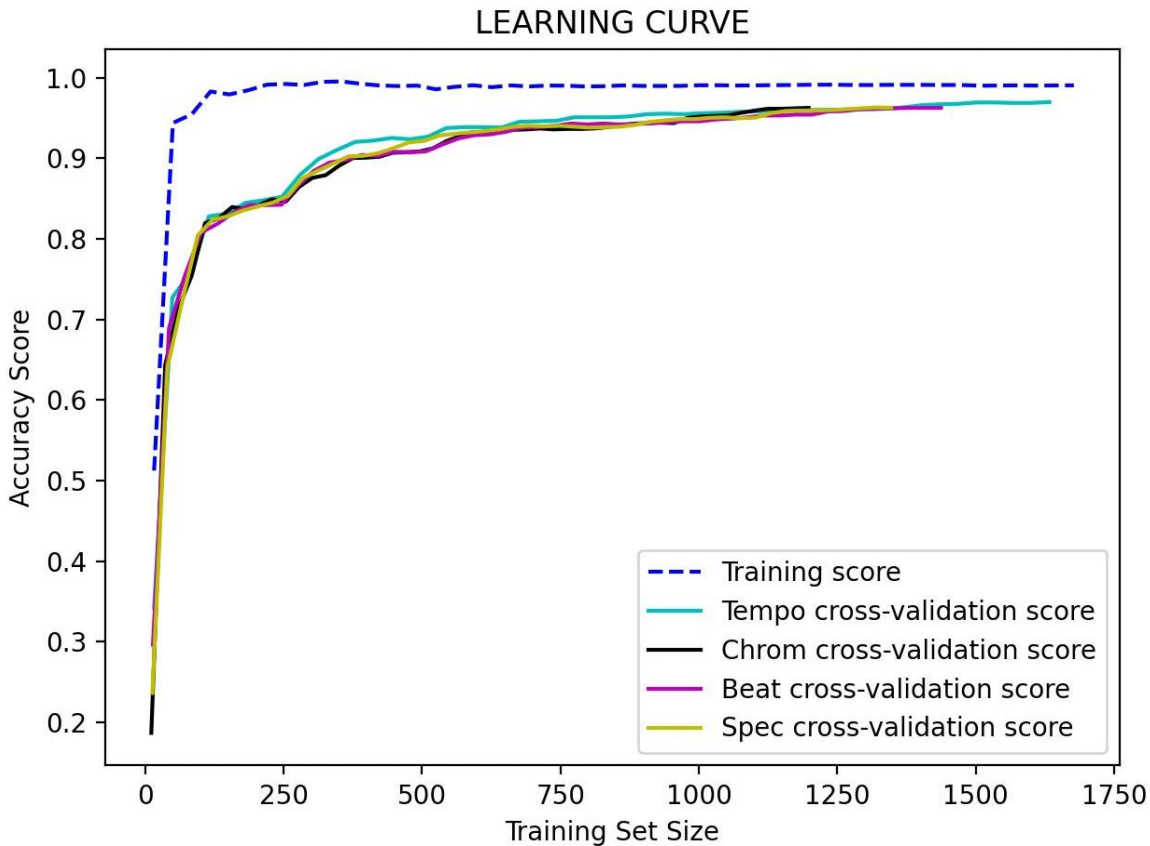


Figure 2. Validation results

Conclusion. Our research has ventured into the realm of hip-hop and R&B music, driven by the quest to refine Beats Per Minute (BPM) detection. Through meticulous analysis, we unveiled that the Tempogram representation emerges as the most accurate, particularly in handling tempo variations—a common characteristic of these genres.

The integration of Convolutional Neural Networks (CNNs) was crucial, resulting in superior BPM detection accuracy and robustness across diverse musical styles. Beyond BPM, our approach unearthed genre-specific insights, promising more nuanced music recommendations.

Our work demonstrates the transformative potential of deep learning in music analysis, inspiring a deeper exploration of rhythmic intricacies. In conclusion, this research not only advances BPM detection but also paves the way for a more profound understanding of the rhythmic diversity that defines hip-hop and R&B music.

Reference list

- [1] E. Parada-Cabaleiro, M. Schmitt, A. Batliner, B. Schuller, and M. Schedl. Automatic recognition of texture in renaissance music. International Society for Music Information Retrieval Conference (ISMIR), 2021.
- [2] D. Guigue and C. de Paiva Santana. The structural function of musical texture: Towards a computer-assisted analysis of orchestration. Journes d'Informatique Musicale (JIM), 2018.

[3] Z.Wang, D.Wang, Y.Zhang, and G.Xia. Learning interpretable representation for controllable polyphonic music generation. International Society for Music Information Retrieval Conference (ISMIR), 2020.

Authors' contribution

Katsiaryna Tushynskaya – led the research on evaluation the quality of BPM detection and analyzing.

Maryia Zyranava – investigated and compared different audio representations, investigated methods of audio features' extracture, developed CNN and analyzed its output.

Aleh Asadchy – prepared a dataset of representations, developed and trained CNN as well as analyzed its output.

УЛУЧШЕНИЕ РАСПОЗНАВАНИЯ BPM В ХИП-ХОПЕ И R&B С ПОМОЩЬЮ АУДИОПРЕДСТАВЛЕНИЙ И КОНВОЛЮЦИОННЫХ НЕЙРОННЫХ СЕТЕЙ

Е.В. Тушинская

*Ассистент кафедры
информатики факультета
компьютерных систем и сетей
БГУИР, студент
магистратуры*

М.М. Зырянова

*Студентка выпускного курса
факультета компьютерных
систем и сетей БГУИР*

О.Э. Осадчий

*Студент выпускного курса
факультета компьютерных
систем и сетей БГУИР*

Аннотация. В данном исследовании представлен метод улучшения обнаружения темпа в хип-хоп и R&B музыке путем интеграции аудио представлений со сверточными нейронными сетями (CNN). Анализируя темпограммы, спектрограммы, хромограммы и битограммы оптимизированные для обработки CNN, разработанный нами подход демонстрирует повышенную точность обнаружения BPM в этих жанрах.

Ключевые слова: Количество ударов в минуту (BPM), сверточная нейронная сеть (CNN), темпограмма, битограмма, ритмограмма, обработка аудио информации.

УДК 621.396 (024)

ЛАЗЕРНЫЙ КОМПЛЕКС ДЛЯ ИЗМЕРЕНИЯ ПАРАМЕТРОВ КОЛЁСНЫХ ПАР ПОДВИЖНОГО СОСТАВА



А.А. Новицкий

Ведущий инженер-конструктор,
ООО «РИФТЭК», магистр
технических наук
a.novitski@riftek.com



Г.М. Шахлевич

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

А.А. Новицкий

Окончил Белорусский государственный университет информатики и радиоэлектроники и магистратуру БГУИР. Область производственных и научных интересов связана с разработкой автоматизированных систем бесконтактных измерений и контроля формы и размеров технических объектов, создания новых методик проектирования лазерных триангуляционных датчиков.

Г.М. Шахлевич

Окончил Белорусский государственный университет и аспирантуру Академии наук БССР. Область научных интересов: физико-химическое материаловедение; новые материалы и структуры изделий микро- и нанoeлектроники, инновационные технологии обработки материалов и производства изделий электронно-оптической техники; научно-методическое и организационное сопровождение учебного процесса.

Аннотация. Рассмотрены вопросы, связанные с разработкой сложных измерительных систем на базе лазерных триангуляционных сканеров, а также применение бесконтактных измерений на железнодорожном транспорте для раннего диагностирования износа колёсных пар и повышения безопасности движения. Описана конструкция основных функциональных модулей и характеристики автоматического лазерного комплекса для измерения геометрических формы и размеров колесных пар подвижного состава в процессе его движения.

Ключевые слова: лазеры, триангуляция, бесконтактные измерения, 3D wheel, лазерные сканеры, измерение параметров колёсных пар, измерительный комплекс.

Введение. Метод триангуляционных измерений становится все более распространенным по мере развития микроэлектроники и лазерной техники. Оптико-электронные приборы для бесконтактного измерения и контроля положения, перемещения, размеров, профиля поверхности изделий, их деформации, вибрации, сортировки, распознавания технологических объектов, измерения уровня жидкостей и сыпучих материалов, использующие принцип оптической триангуляции, называются триангуляционными датчиками.

Принцип триангуляционных измерений. Триангуляция – это определение взаимного расположения точек на поверхности при помощи построения сети треугольников. Метод позволяет измерять как относительное изменение расстояния от датчика до контролируемого объекта, так и абсолютную его величину [1]. Контролируемое расстояние может иметь масштаб от долей микрометра до тысяч метров. Расстояния до шероховатых поверхностей измеряются путем использования лазерного

луча в виде узкой линии. Световое пятно, появляющееся на поверхности объекта, захватывается детектором системы и определяется его позиция [2]. Расстояние вычисляется по изменению угла, под которым она находится. Возможные разрешение и точность в основном зависят от расстояния d . Вблизи датчика может быть получено большое изменение угла a_1 , тогда как большие значения приводят к меньшему углу a_2 , что уменьшает точность [3] (рисунок 1).

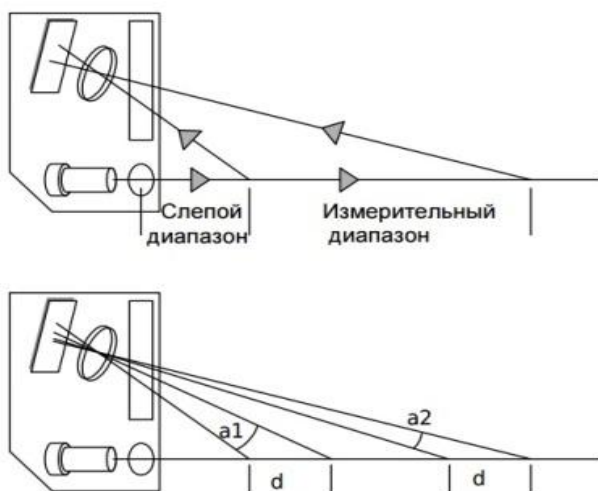


Рисунок 1. Функциональная схема триангуляционного датчика

2D лазерные сканеры. На базе лазерных датчиков построена серия 2D лазерных сканеров. Излучение полупроводникового лазера формируется в виде линии и проецируется на объект. Рассеянное на объекте излучение собирается объективом и передается на двумерную светочувствительную CMOS-матрицу. Полученное изображение контура объекта анализируется сигнальным процессором, который рассчитывает расстояние до объекта (координата Z) для каждой из множества точек вдоль лазерной линии на объекте (координата X). Сканеры характеризуются положением начала рабочего диапазона сканирования (SMR) по координате Z , рабочим диапазоном (MR) по координате Z , рабочим диапазоном по координате X в начале рабочего диапазона по Z (X_{smr}) и в конце рабочего диапазона по Z (X_{emr}) [4] (рисунок 2).

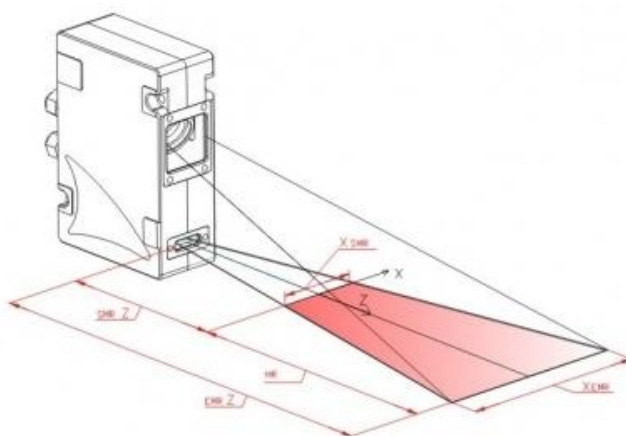


Рисунок 2. Внешний вид 2D-сканера и схема измерений

Автоматическая система контроля колесных пар. На основе лазерных сканеров, была разработана система для бесконтактных измерений геометрических формы и размеров колесных пар подвижного состава в процессе его движения. Система предназначена для бесконтактного автоматического измерения геометрических параметров колесных пар железнодорожного подвижного состава (локомотивов, вагонов, метро, трамваев) и использует комбинацию 2D лазерных сканеров (5 или 7 шт.), установленных по обе стороны рельса и откалиброванных в одной системе координат (рисунки 3, 4, 5).

Цикл измерения начинается, когда индуктивный датчик обнаруживает колесо. Профили колес снимаются в момент прохождения их через зону сканирования. Результаты измерений от каждого датчика посылаются через *Ethernet* на управляющий компьютер для обработки данных, построения профилей колес и расчета размеров и сохраняются в базе данных.

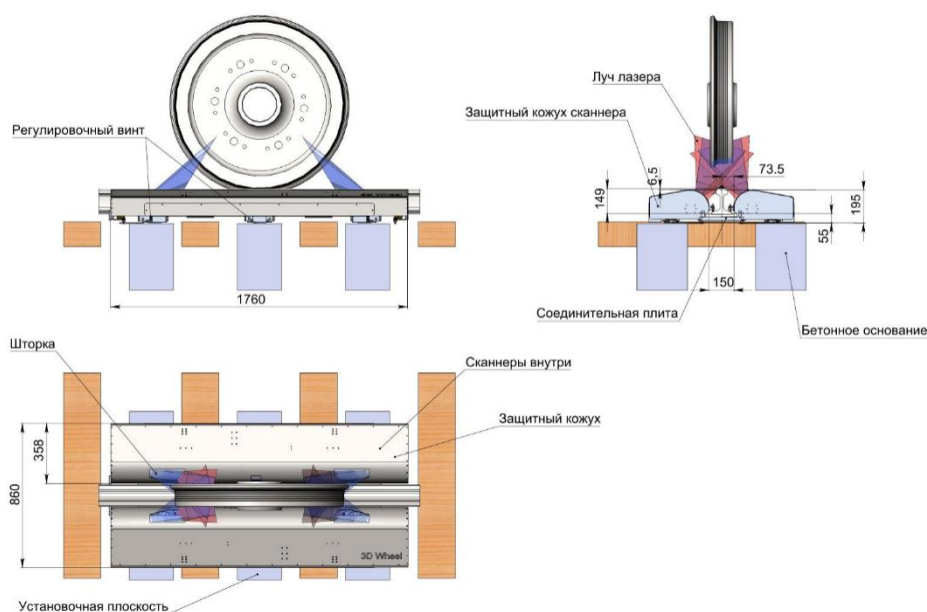


Рисунок 3. Схема измерительной системы

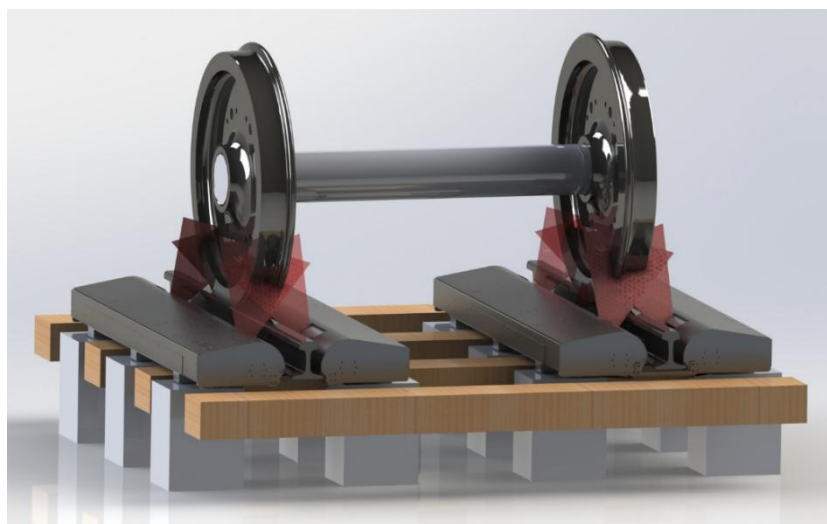


Рисунок 4. Внешний вид измерительной системы

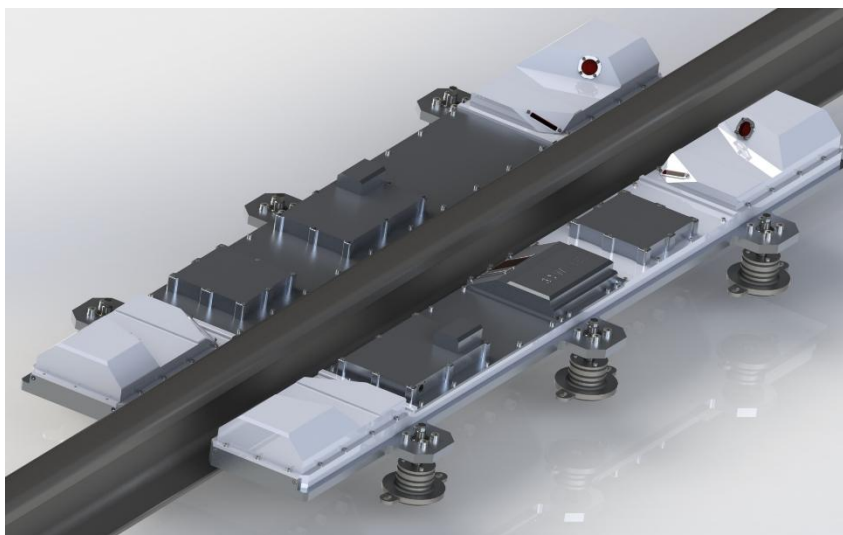


Рисунок 5. Внешний вид измерительной системы без защитных кожухов

Измерительная система позволяет производить измерения с достаточно высокой точностью. Измеряемые параметры и погрешности измерения приведены в таблице 1.

Таблица 1. Параметры, измеряемые системой контроля колёсных пар на ходу

Параметр	Погрешность
Полный профиль колеса	$\pm 0,1$ мм
Высота гребня	$\pm 0,1$ мм
Ширина гребня	$\pm 0,1$ мм
Крутизна гребня	$\pm 0,1$ мм
Толщина бандажа	$\pm 0,1$ мм
Ширина бандажа	$\pm 0,1$ мм
Межбандажное расстояние	$\pm 0,05$ мм
Диаметр колеса	$\pm 0,2$ мм

Заключение. Автоматическая система контроля колесных пар на основе лазерных сканеров позволяет выполнять бесконтактные измерения геометрической формы и размеров колесных пар подвижного состава в процессе его движения. По точности она не уступает традиционным инструментальным измерениям, но значительно превосходит их по производительности и информативности.

Список литературы

- [1] Игнатов, А.Н. Оптоэлектронные приборы и устройства. Москва: Экотрендз, 2006. – 258 с. .
- [2] Меркишин Г.В. Многооконные оптико-электронные датчики линейных размеров. М.: Радио и связь. - 1986. – 168 с.
- [3] Апенко М.М., Дубовик А.С., Дурейко Г.В. Прикладная оптика: учеб. пособие М.: Машиностроение, 1992.– 470 с
- [4] Новицкий А.А., Шахлевич Г.М., Романов А.В. Оптико-электронные приборы для прецизионных измерений формы и размеров изделий / Приборостроение –2016: Материалы 9-ой Междунар. научно-технической конференции, Минск 16-18 ноября 2016 г., БНТУ: Минск, 2016. – с. 352-353.

[5]. Новицкий А.А., Шахлевич Г.М. Устройство лазерного и визуального контроля внутренней поверхности отверстий. Новые направления развития приборостроения: материалы 12-й Междунар. научно-технической конференции молодых ученых и студентов, Минск, 17–19 апреля 2019 г., БНТУ. – Минск, 2019. – с.68-69.

Авторский вклад

Новицкий Александр Александрович – разработка структурных и принципиальных схем электронных модулей системы, проектирование конструкции измерительного комплекса, контроль процесса изготовления его узлов, руководство процессом монтажа и сборки, настройка, регулировка и испытания комплекса, представление материалов для написания статьи.

Шахлевич Григорий Михайлович – постановка задачи и руководство процессом проектирования, участие в разработке структурных и функциональных схем, выборе оптоэлектронных датчиков, согласовании аппаратной и программной частей комплекса, обработка результатов испытаний, написание окончательного варианта статьи.

LASER COMPLEX FOR MEASUREMENT OF PARAMETERS ROLLING STOCK WHEEL PAIRS

A.A. Novitski

*Lead Design Engineer,
LLC "RIFTEK," Master*

1. technical sciences

2.

G.M. Shakhlevich

*Associate Professor, Department of Electronic
Engineering and technologies of BGUIR, PhD of
physical and mathematical sciences, associate
professor*

Abstract. Issues related to the development of complex measuring systems based on laser triangulation scanners, as well as the use of contactless measurements on railway transport for early diagnosis of wear of wheelsets and improving traffic safety are considered. Described is design of main functional modules and characteristics of automatic laser complex for measurement of geometrical shape and dimensions of wheelsets of rolling stock during its movement.

Keywords: lasers, triangulation, contactless measurements, 3D wheel, laser scanners, measurement of wheel pair parameters, measuring complex.

УДК 004.4:004.67

МОДЕЛИРОВАНИЕ СИГНАЛОВ В ПРОГРАММНОЙ СРЕДЕ LABVIEW С ВОЗМОЖНОСТЬЮ ЭКСПОРТА И ОБРАБОТКИ ДАННЫХ В MS EXCEL



Г.А. Пискун

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

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. Представлен процесс моделирования генератора сигналов в программной среде для разработки виртуальных приборов *LabVIEW*. Описан принцип генерации сигналов, а также функция добавления к нему шума. Рассмотрена возможность экспорта данных о выходном сигнале в *MS Excel*.

Статья подготовлена в рамках выполнения гранта Президента Республики Беларусь на 2024 год в сфере образования (распоряжение Президента Республики Беларусь от 3 января 2024 г. № 4рп «О предоставлении грантов Президента Республики Беларусь на 2024 год»).

Ключевые слова: генератор сигналов, шум, данные

Введение. В большинстве случаев, полученные в результате проведения экспериментов данные без предварительной обработки обладают малой полезной информативностью. В связи с этим, очень важно использовать инженерное программное обеспечение, в котором уже заложен обширный функциональный инструментарий, позволяющий выполнить сбор и обработку данных.

Одной из наиболее известных сред в которой заложены различные функции работы с экспериментальными данными является *LabVIEW*. Также она обладает широким набором возможностей для отладки сложных программ, тестирования реальных систем измерения, регулирования и развернутого анализа получаемых данных [1].

Рассмотрим встроенную в данную систему функцию для генерации сигналов и шумов, которая используется при формировании детерминированных и случайных сигналов с заданным набором параметров.

Программная реализация работы с сигналами в *LabVIEW*. Генерация сигналов является важным инструментом в научных и технических приложениях, обеспечивая контролируемую среду для исследований в различных областях.

В *LabVIEW* реализована возможность работы с сигналами, в частности, есть встроенные виртуальные приборы (ВП), которые представляют многофункциональные

программно-регулируемые генераторы сигналов с широким набором контролируемых параметров; элементы, предназначенные для генерации наиболее широко применяемых детерминированных периодических сигналов, а также для генерации шумов с различными законами амплитудного и спектрального распределения.

Есть функция анализа сигналов при помощи генераторов с заданной длительностью сигналов; гармоническими колебаниями и шумом; отрезками синусоидального, импульсного, пилообразного, $\sin(x)/y$, прямоугольного и частотно-модулированного сигналов; синусоидальными, треугольными, прямоугольными, пилообразными и произвольными колебаниями любой длительности; равномерными, гауссовскими, периодическими случайными шумами и двоичной последовательностью максимальной длины; гамма-шумом, пуассоновскими, биномиальными шумами, шумом Бернулли [1 – 3].

Дополнительной возможностью является анализ шумов, которые являются чисто случайными сигналами, мгновенное значение которых или фазу невозможно предсказать во времени. Они ухудшают качество полезного сигнала, могут генерироваться внутри операционных усилителей или связанных с ним пассивных элементов, либо попадать в схему из внешних источников [4, 5].

Принцип создания генератора сигналов. В «*Functions palette*» на «*Block Diagram*» необходимо выбрать «*Signal Generation by Duration*»: «*Function*» → «*Signal Processing*» → «*Signal Generation*» → «*Signal Generation by Duration*». Для получения информации по какому-либо виртуальному прибору (ВП) требуется активизировать его изображение при помощи левой кнопки мыши и выбрать пункт меню «*Help*» (рисунок 1, а).

Рассматриваемый ВП имеет девять входов и три выхода. Входы: сбросить фазу, длительность, тип сигнала, число выборок, частота, амплитуда, постоянная составляющая, вход фазы, заполнение цикла прямоугольного колебания (%). Выходы: сигнал, частота выборок, выход фазы (рисунок 1, б).

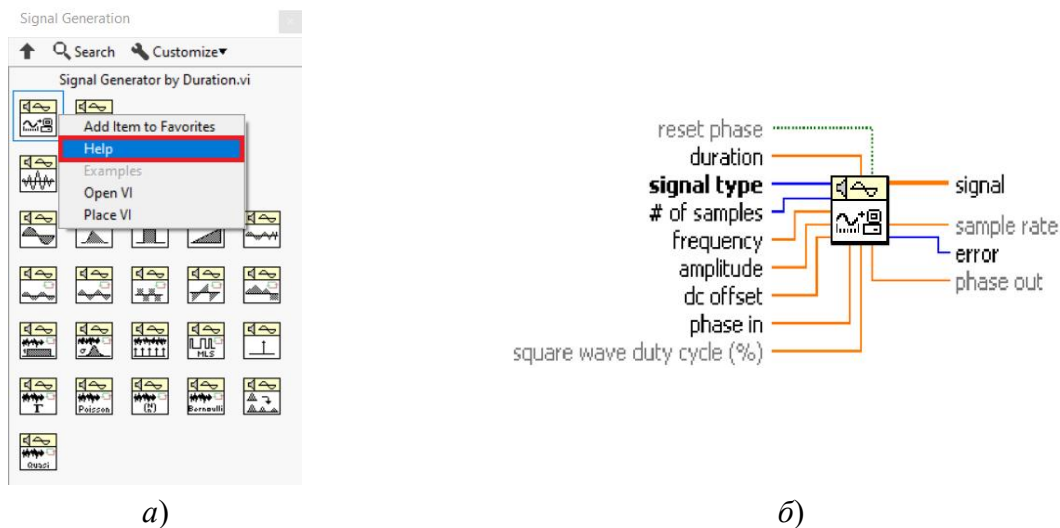


Рисунок 1 – Выбор генератора сигналов с заданной длительностью в среде LabVIEW (а) и его обозначение (б)

ВП генерирует «*signal*», имеющий форму, задаваемую на входе «*signal type*». Вход «*reset phase*» определяет начальную фазу выходного сигнала. По умолчанию на входе установлено состояние «ИСТИНА». При этом начальная фаза сигнала устанавливается в соответствии со значением на входе «*phase in*». Если на входе «*reset phase*» установлено состояние «ЛОЖЬ», то начальная фаза устанавливается равной значению фазы на «*phase out*» при последнем выполнении этого ВП.

Далее добавим структуру «*While Loop*» («*Functions*» → «*Structures*» → «*While Loop*») (рисунок 2, а), перенесем туда «*Signal Generation by Duration*», в которой добавим все необходимые индикаторы и элементы управления (рисунок 2, б).

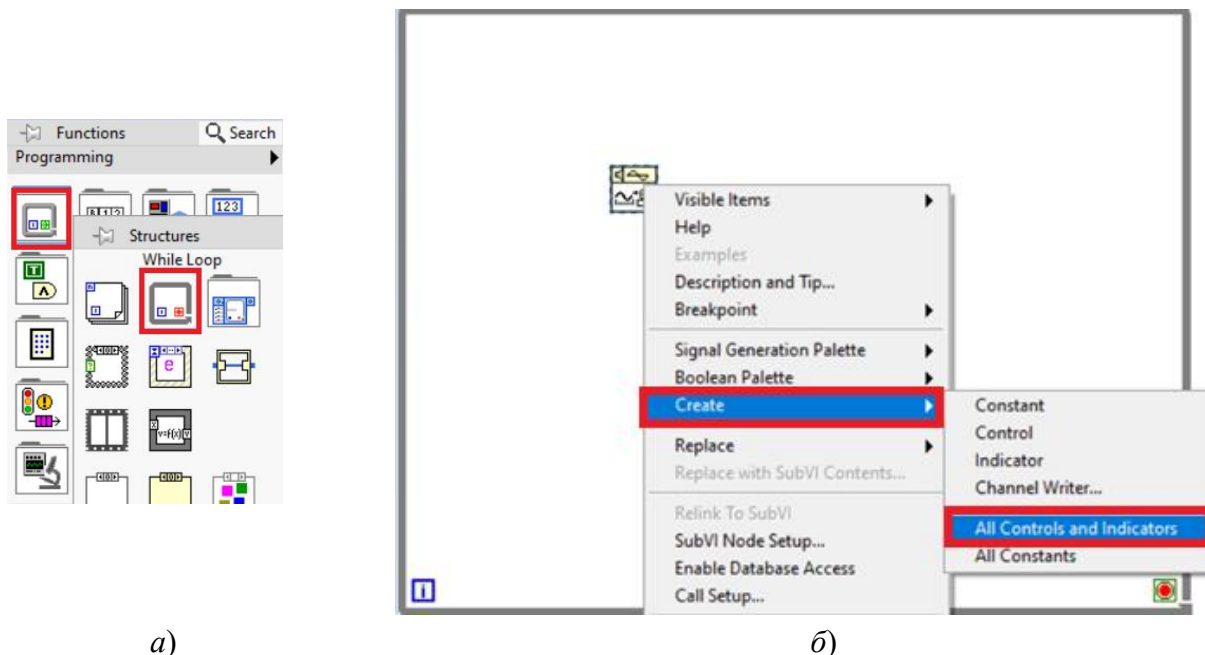


Рисунок 2 – Выбор структуры «*While Loop*» (а) и добавление необходимых элементов управления и индикаторов для «*Signal Generation by Duration*» (б)

Добавим «*Waveform Graph*» на «*Front Panel*» посредством открытия «*Controls*» → «*Graph*» → «*Waveform Graph*» (рисунок 3).

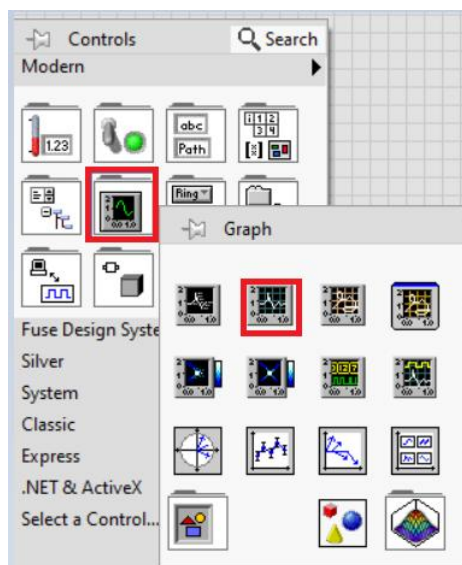


Рисунок 3 – Выбор «*Waveform Graph*»

На «*Block Diagram*» соединим выход «*signal*» элемента «*Generation by Duration*» и вход «*Waveform Graph*». Для этого предварительно удалим индикатор «*signal*» с помощью выделения и последующего нажатия кнопки «*Delete*», после чего связываем выход «*signal*» с входом «*Waveform Graph*» напрямую (рисунок 4).

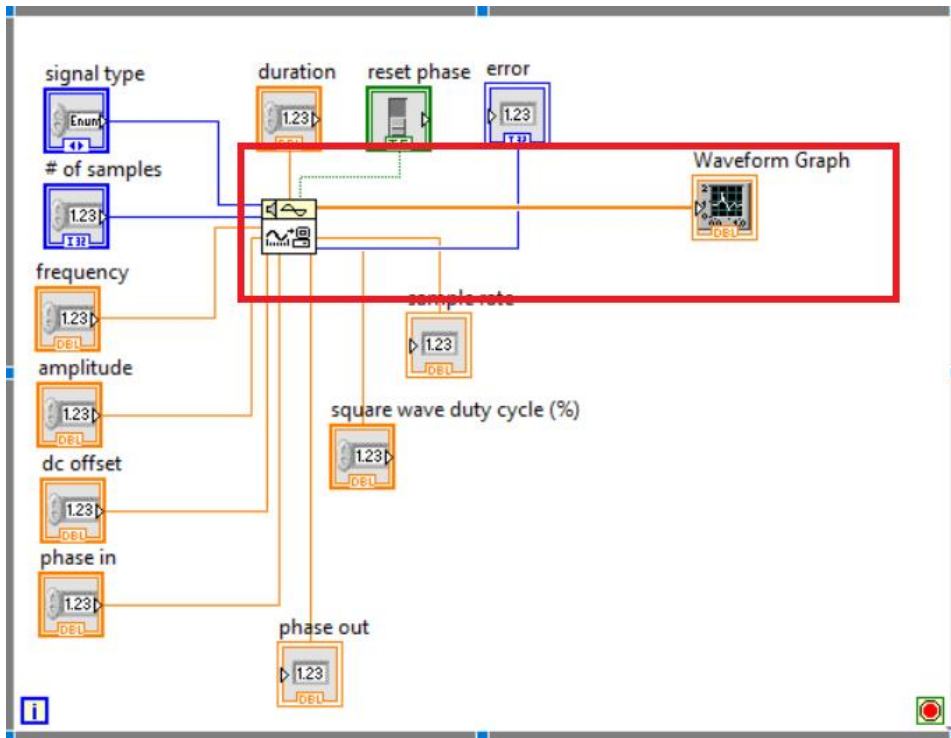


Рисунок 4 – Обозначение соединения выхода «*signal*» элемента «*Generation by Duration*» и входа «*Waveform Graph*»

Добавим кнопку для виртуального завершения цикла «*Loop Condition*» → «*Create Control*» (рисунок 5).

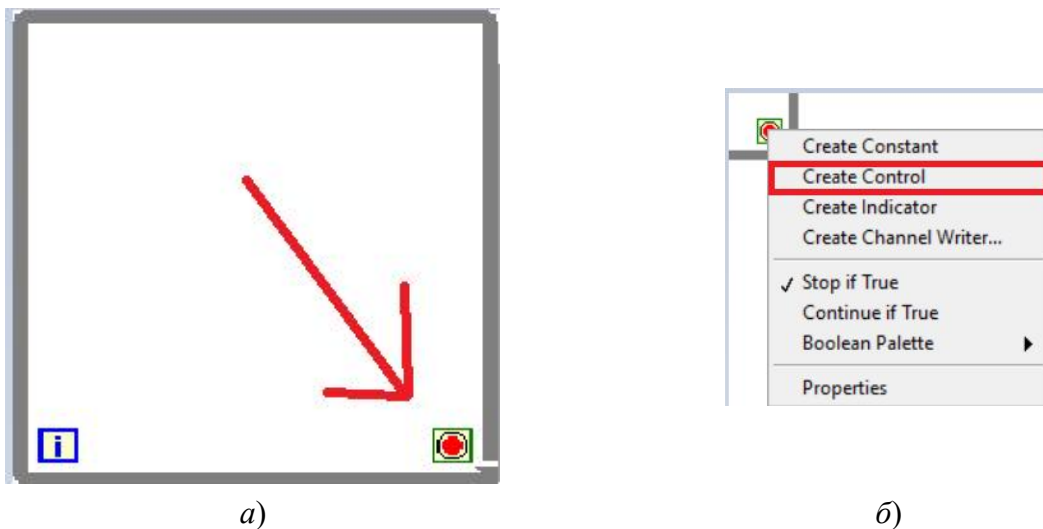


Рисунок 5 – Расположение кнопки завершения цикла «*Loop Condition*» (а) и ее добавление (б)

Далее разместим на «*Front Panel*» элементы управления и индикаторы как удобно, заполняем поля: тип сигнала («*signal type*»), длительность («*duration*»), число выборок («*# of samples*»), частота («*frequency*»), амплитуда («*amplitude*»), постоянное смещение («*dc offset*»), начальное значение фазы («*phase in*») и запускаем программу (рисунок 6).

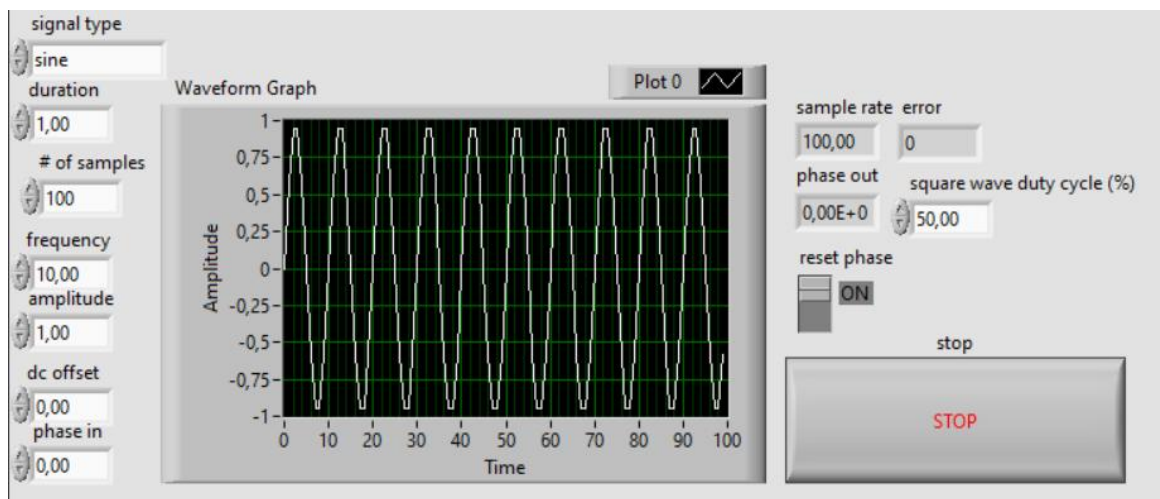
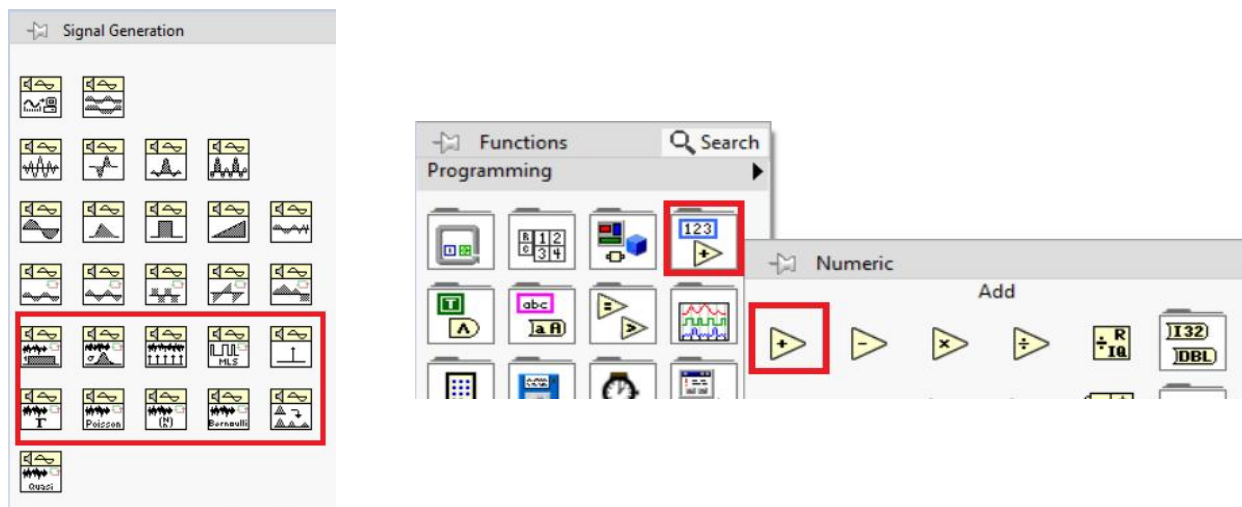


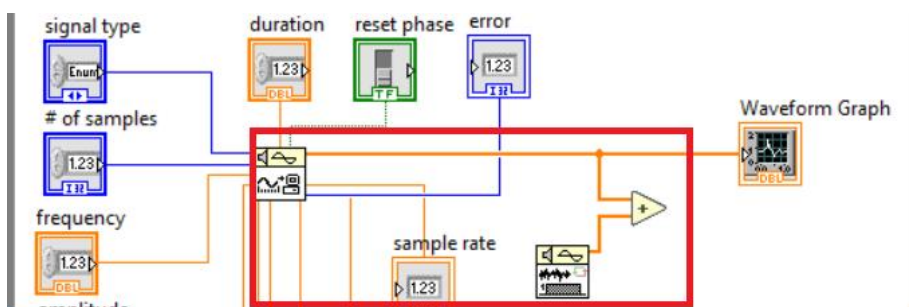
Рисунок 6 – Пример «Front Panel» с расположенными элементами управления, индикаторами и осциллографом, отображающим синусоидальный сигнал

Добавление к сигналу шума. Для этого необходимо на «Block Diagram», из панели «Sig Generation» выбрать генератор шума и внести его на блок диаграмму, после добавить к нашему сигналу шум с помощью функции «Function» → «Numeric» → «Add» (рисунки 7).



а)

б)



в)

Рисунок 7 – Выбор генератора шума (а), выбор генератора «Numeric» (б) и добавление шума к сигналу (в)

После добавим элементы управления амплитудой шума с помощью наведения курсора мыши на вход «*amplitude*» и нажатия правой кнопки мыши, что позволит вызвать контекстное меню. В меню выбираем «*Create*» → «*Control*» (рисунок 8).

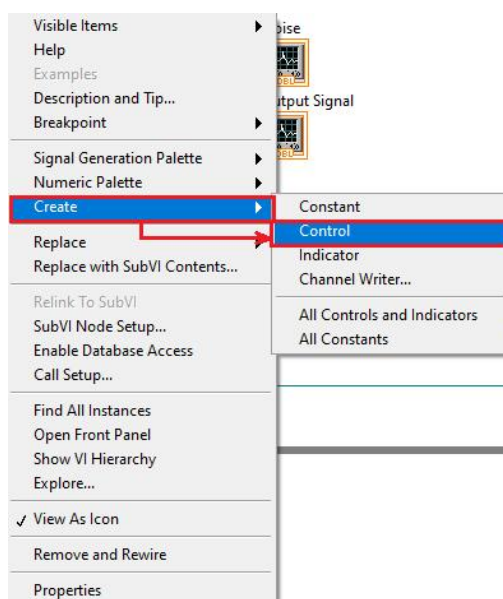


Рисунок 8 – Добавление «*Control*» на «*Block Diagram*»

Добавим на «*Block Diagram*» функцию для отображения спектра измененного сигнала. Для этого выберем функцию «*Auto Power Spectrum*» («*Function*» → «*Signal Processing*» → «*Spectral*» → «*Auto Power Spectrum*») (рисунок 9).

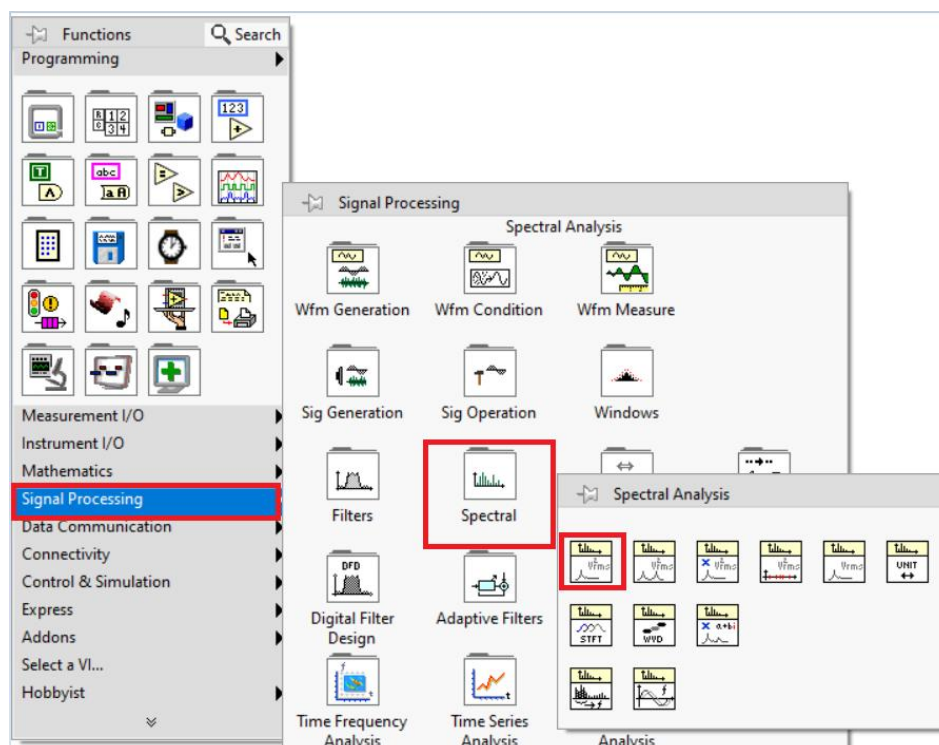


Рисунок 9 – Выбор функции «*Auto Power Spectrum*»

Применим эту функцию к выходному сигналу (рисунок 10).

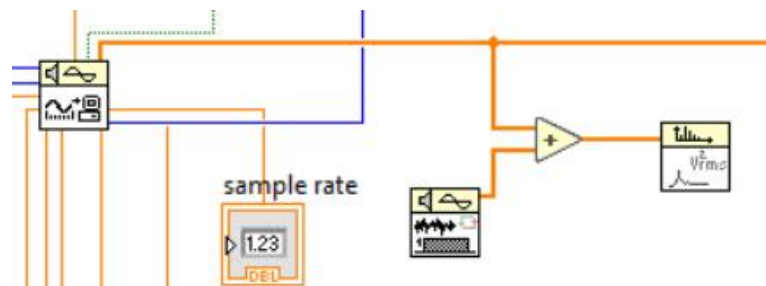


Рисунок 10 – Фрагмент диалогового окна на котором отображено применение функции «Auto Power Spectrum» к выходному сигналу

Для одновременного отображения шума, выходного сигнала и спектра выходного сигнала необходимо добавить еще три «Waveform Graph» (рисунок 11 и 12).

Стоит отметить, что элементы соединяются проводниками различных цветов и толщин. Это обусловлено тем, что по проводнику синего цвета передаются данные целого типа («integer»), по оранжевому – вещественного («real», «float»), по зеленому – логического («boolean»). Разность толщин показывает, что по самой тонкой передается переменная, чуть толще – массив, еще более толстая – структура [2 – 5].

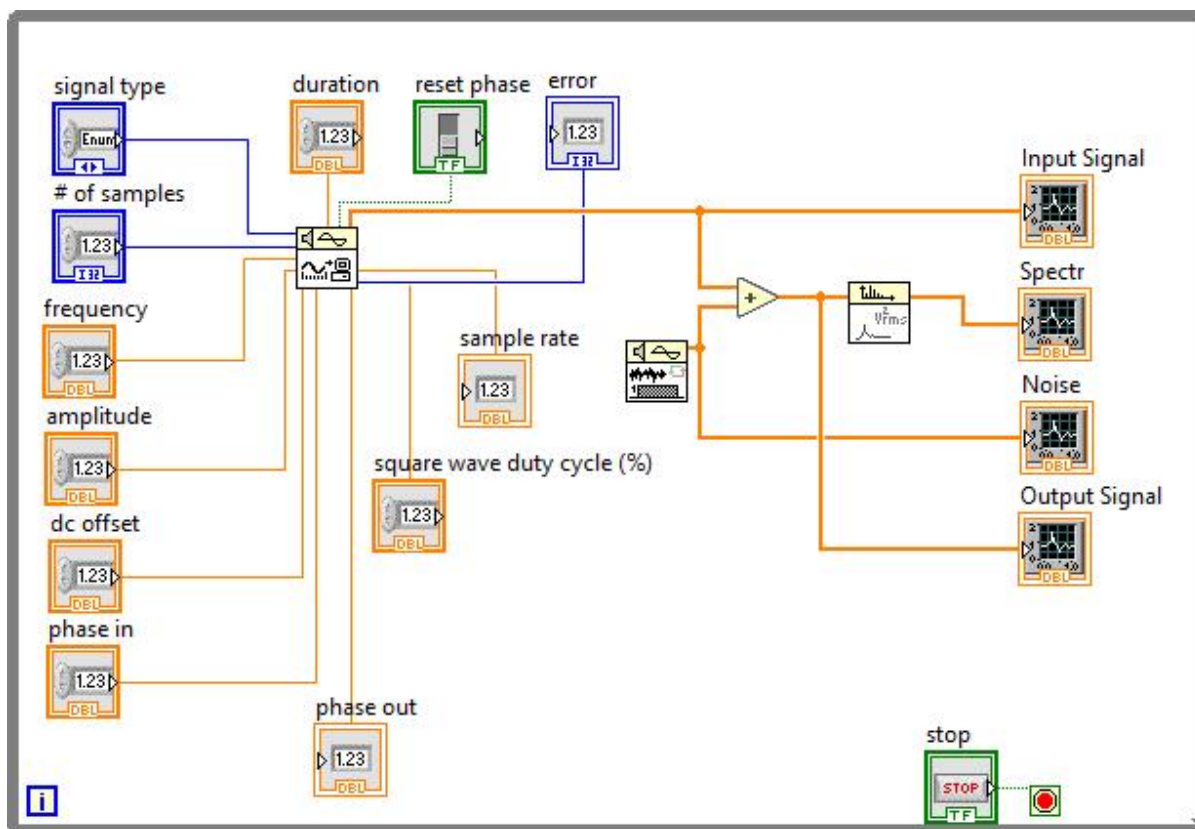


Рисунок 11 – Диалоговое окно «Block Diagram» с отображением виртуального прибора генератора сигналов

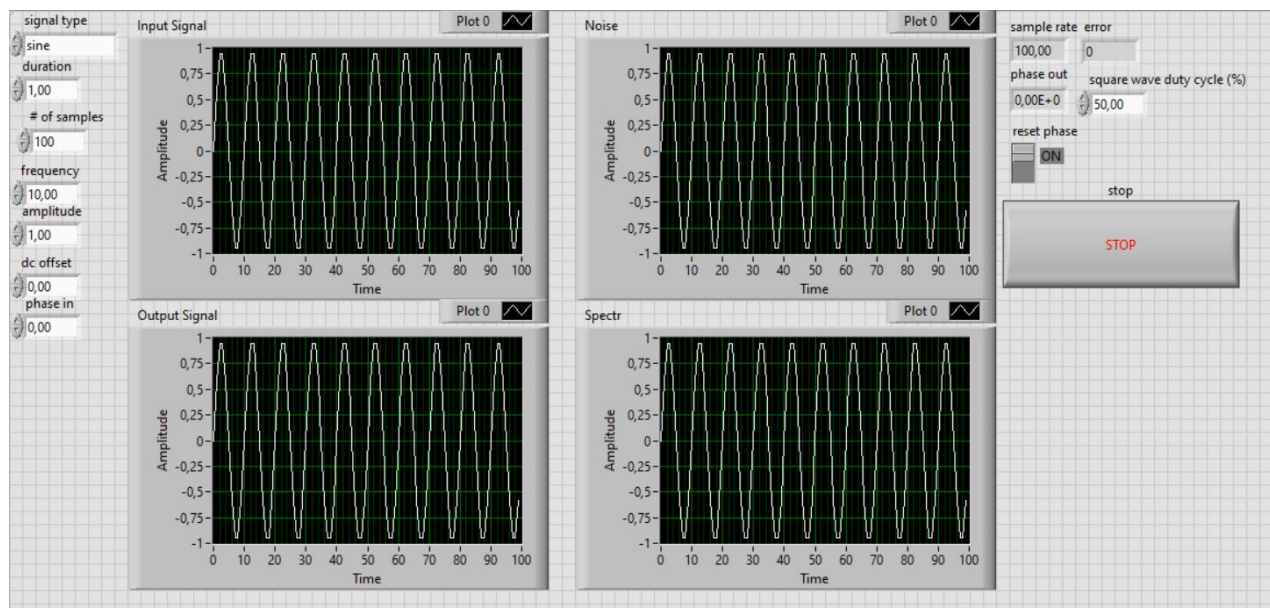


Рисунок 12 – «Front Panel» с отображением осциллограмм синусоидального сигнала, шума, выходного сигнала и спектра выходного сигнала

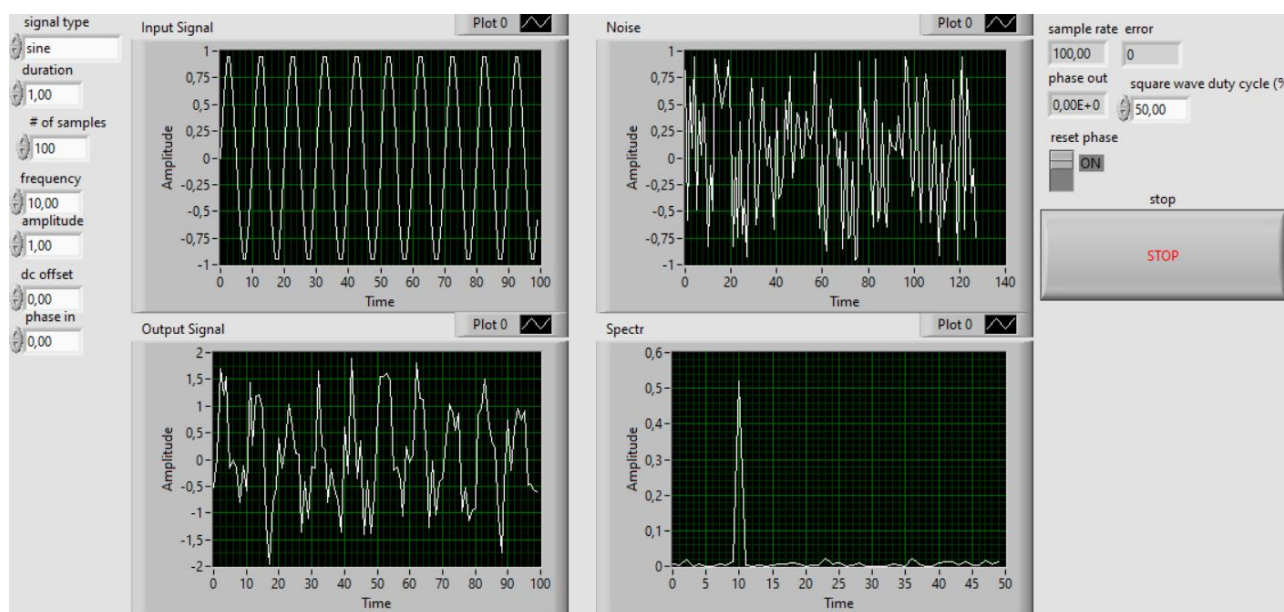


Рисунок 13 – Запуск генератора сигналов с наложенным шумом

Экспорт данных выходного сигнала в MS Excel. Одной из функций среды *LabVIEW* является реализация экспорта данных в такую программу для работы с электронными таблицами и визуализацией результатов экспериментов, как *MS Excel*.

MS Excel позволяет легко отслеживать и анализировать большие объемы информации, создавать графики и диаграммы, необходимые для наглядного представления данных и принятия решений на основе проведенного анализа.

Для экспорта данных необходимо щелчком правой кнопкой мыши по «*Waveform Graph*», на котором отображается выходной сигнала, выбрать «*Export*» → «*Export Data To Excel*» (рисунок 14).

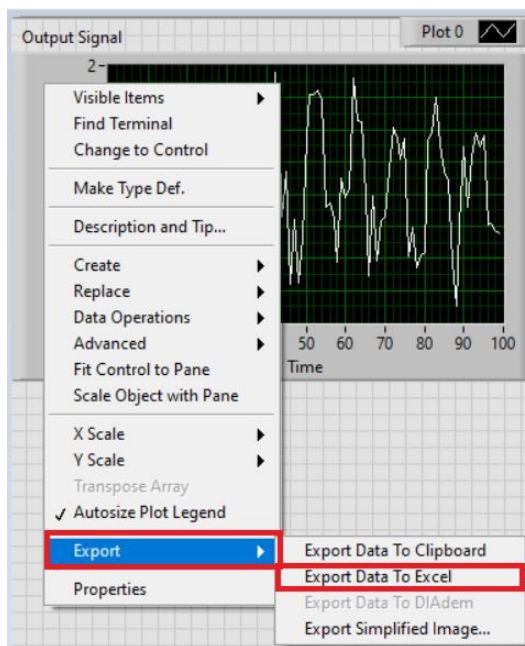


Рисунок 14 – Экспорт данных в MS Excel

С помощью экспортированных данных можно построить график и описать свойства полученного сигнала. Для этого выделим все экспортированные данные, далее выберем «Вставка» → «Вставить точечную (X, Y)» или пузырьковую диаграмму → «Точечная с гладкими кривыми и маркерами» (рисунок 15).

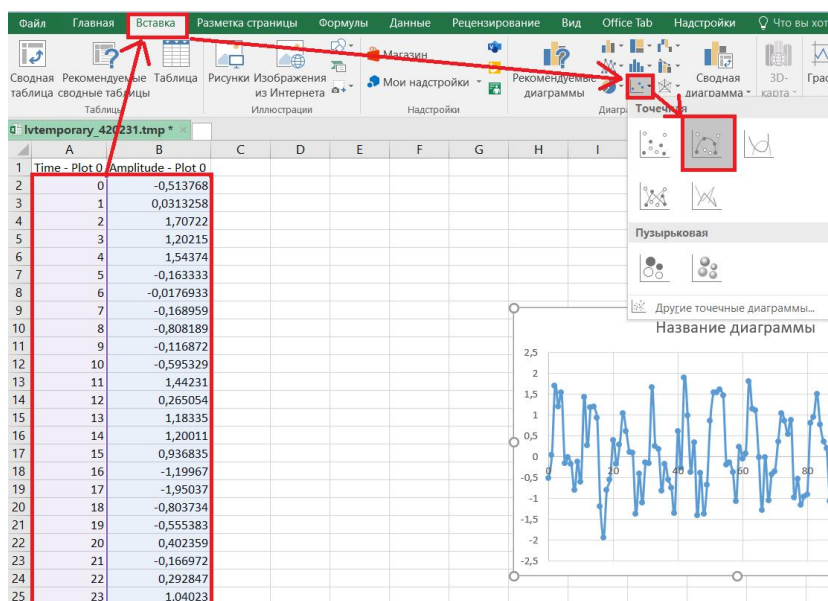


Рисунок 15 – Построение графика выходного сигнала по экспортированным данным

Заключение. В работе подробно рассмотрены возможности программной среды LabVIEW, необходимые для генерации сигналов и шумов. Приведен принцип разработки генератора сигналов, рассмотрена функция добавления к сигналу шума, показаны осциллограммы синусоидального сигнала, шума, выходного сигнала и спектра выходного сигнала. Рассмотрена возможность экспорта данных о выходном сигнале в MS Excel.

Список литературы

- [1] Автоматизация физических исследований и эксперимента: компьютерные измерения и виртуальные приборы на основе *LabVIEW 7* (30 лекций) / П.А. Бутыркин [и др.]. – Москва : ДМК Пресс, 2005. – 264 с.
- [2] Васильев, А.С. Основы программирования в среде *LabVIEW* / А.С. Васильев, О.Ю. Лашманов. – Санкт-Петербург : Университет ИТМО, 2015. – 82 с.
- [3] Евдокимов, Ю.К. *LabVIEW* для радиоинженера: от виртуальной модели до реального прибора / Ю.К. Евдокимов, В.Р. Линдваль, Г.И. Щербаков. – Москва : ДМК Пресс, 2007. – 400 с.
- [4] Евдокимов, Ю.К. *LabVIEW* в научных исследованиях / Ю.К. Евдокимов, В.Р. Линдваль, Г.И. Щербаков. – Москва : ДМК Пресс, 2018. – 400 с.
- [5] *LabVIEW: Практикум по основам измерительных технологий* / В.К. Батоврин [и др.]. – Москва : ДМК Пресс, 2005. – 208 с.

SIGNAL MODELING IN LABVIEW SOFTWARE ENVIRONMENT WITH THE POSSIBILITY OF EXPORTING AND PROCESSING DATA IN MS EXCEL

G.A. Piskun

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

Abstract. The process of modeling a signal generator in the software environment for developing virtual instruments LabVIEW is presented. The principle of generating signals, as well as the function of adding noise to it, is described. The possibility of exporting output signal data to MS Excel has been considered.

The article was prepared as part of the implementation of the grant of the President of the Republic of Belarus for 2024 in the field of education (Order of the President of the Republic of Belarus dated January 3, 2024 No. 4пр «On the provision of grants of the President of the Republic of Belarus for 2024»).

Keywords: signal generator, noise, data

УДК 004.021:004.75

ВНЕДРЕНИЕ ИНТЕГРАЦИИ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ В СТРОИТЕЛЬСТВЕ



С.А. Мигалевич
Магистр технических наук,
начальник центра
информатизации и
инновационных разработок
migalovich@bsuir.by



А.Н. Марков
Магистр технических
наук, заместитель
начальника центра
информатизации и
инновационных
разработок
a.n.markov@bsuir.by



Г.А. Богрецов
Студент Белорусского
государственного
университета
информатики и
радиоэлектроники
herman.bog@mail.ru

С.А. Мигалевич

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов: вибродиагностика, разработка метода вейвлет-анализа изделий машиностроения.

А.Н. Марков

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов: вычислительные системы, облачные вычисления (CLOUD COMPUTING), распределенные вычислительные системы, балансировка нагрузки вычислительных систем (load balancing).

Г.А. Богрецов

Студент Белорусского государственного университета информатики и радиоэлектроники факультета компьютерного проектирования.

Аннотация. Целью данного исследования было продемонстрировать потенциал современных информационных технологий в области строительства, проектирования и эксплуатации строительных объектов. Дополнительной задачей было рассмотрение возможности комбинирования или совместного использования различных информационных технологий, таких как технология больших данных и BIM-технология, для решения задач в строительстве. Кроме того, исследование ставило целью выявить потенциал повышения эффективности определенных строительных процессов за счет применения таких комбинированных информационных технологий.

Ключевые слова: большие данные, строительные процессы, отслеживание рисков, диспетчеризация, автоматизация эксплуатации.

Введение. Применение информационно-технологических решений в процессе проектирования и строительства, особенно обработка больших объемов данных, является крайне актуальным с точки зрения сокращения расходов и увеличения производительности. Для крупных и затратных проектов это приносит значительные преимущества. Кроме того, автоматизация процессов помогает снизить риск человеческого вмешательства, уровень ошибок и другие негативные факторы. Помимо этого, использование больших объемов данных оказывает положительное влияние на

логистику, прогнозирование, планирование, эффективное распределение ресурсов и многое другое [1].

Управление данными в проектах строительства. Строительная отрасль генерирует огромное количество данных на различных этапах проекта, от проектирования до эксплуатации. Эти данные, как структурированные, так и неструктурированные, часто распределены и хранятся в цифровом формате. Недостаток систем для управления и использования этими данными приводит к трудностям в их классификации, сохранении и обновлении. Отсутствие централизованной базы данных осложняет управление проектами и препятствует инновациям.

В сравнении с другими отраслями, строительство пока не имеет эффективных баз данных на уровне предприятия и проекта. Традиционные методы управления качеством в строительстве становятся менее эффективными в эпоху больших данных. В то время как большие данные открывают новые возможности для оптимизации управления качеством, их недостаточное использование приводит к потере ресурсов и замедляет развитие отрасли.

С учетом разброса данных и отсутствия систематического подхода к их управлению, необходимо разработать эффективные методы обработки, анализа и использования данных в строительстве. Это позволит сократить ручную работу, повысить эффективность управления проектами и принимать более обоснованные решения в реальном времени.

Платформа интеграции данных BIM. Моделирование информации о зданиях (BIM) основано на различных связанных данных, генерируемых в ходе реализации строительного проекта. BIM создает базу данных с информацией, собранной на протяжении всего жизненного цикла строительного проекта, и разрушает единичные связи между участниками проекта. Модель меняет пассивное положение, в котором традиционные проекты полагаются на бумажные материалы или персонализированную сетевую коммуникацию для передачи информации о проекте, позволяя участникам в реальном времени понимать прогресс проекта и использовать интернет-технологии для поиска самых свежих, точных и полных данных о проекте. Это снижает проблемы качества, вызванные низкой эффективностью сотрудничества, и является важным способом реализации уточнения и управления информацией в строительной отрасли.

Появление платформы интеграции данных BIM принесло большие преимущества для непрерывности и согласованности данных. Управление проектом на всех этапах жизненного цикла основано на трехмерной модели [2]. Каждый участник непрерывно вводит и обновляет модель BIM и извлекает базовую информацию, такую как геометрические параметры, физические характеристики компонентов, а также информацию о таких управляющих факторах, как качество проекта, безопасность, стоимость, график и цивилизованное строительство, используются в качестве расширенной информации об атрибутах компонентов для реализации бизнес-процесса управления качеством продукции проекта [3]. Интеграция информации, интеграция информации на всем жизненном цикле и интеграция информации организации управления привели к полной информационной модели. Вся база данных BIM на протяжении всего жизненного цикла обновляется в реальном времени. Каждый участник может делиться данными из разных ракурсов в пределах своей компетенции и работать в кооперативном режиме. Изменение способа обмена информацией BIM показано на рисунке 1.

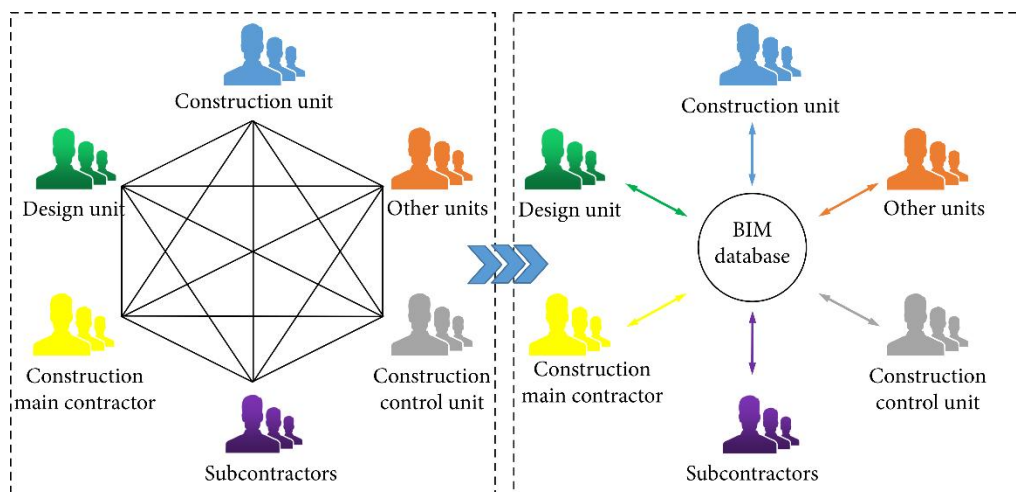


Рисунок 1. Изменения в режиме обмена информацией

VIM: исходный код для больших данных в строительстве. *Big data* – это набор данных, которые невозможно управлять и обрабатывать с использованием обычных программных инструментов в определенные временные рамки. Это огромное количество данных, требующее новой модели обработки для более глубокого принятия решений, понимания и оптимизации процессов. *Big data* – это разнообразный информационный ресурс с высоким темпом роста. Основа технологии больших данных заключается в специализированной обработке данных, которая ищет информацию и добавляет ценность данным путем увеличения способности обработки данных. *VIM* имеет мощную систему хранения данных, включая слои данных, модельный слой и слой приложений информации, которые создают эффективную платформу для интеграции информации. На основе информационных данных о строительном проекте модель определяет основные данные, такие как атрибуты сбора, физические структурные атрибуты и функциональные атрибуты компонентов, и создает трехмерную модель информации о здании на основе этих данных. *VIM* может реализовать динамичное, интегрированное и визуальное управление информацией. Объекты модели связаны с атрибутивной информацией и данными отчетности. Внесение, изменение, удаление и обновление атрибутивной информации объектов модели приведет к мгновенному обновлению данных отчетности, связанных с ними, что обеспечит динамичную передачу информации. С развитием проекта и глубоким применением трехмерных моделей зданий информация о фазах проектирования, подготовки к строительству, строительства, завершения и эксплуатации строительного проекта непрерывно интегрируется на основе трехмерной модели для обеспечения непрерывности и согласованности информации о фазах, в конечном итоге формируя *VIM*-модель информации о продукте проекта и бизнес-процесса, информации о жизненном цикле и информации об управлении организацией. Источник данных *VIM* показан на рисунке 2.

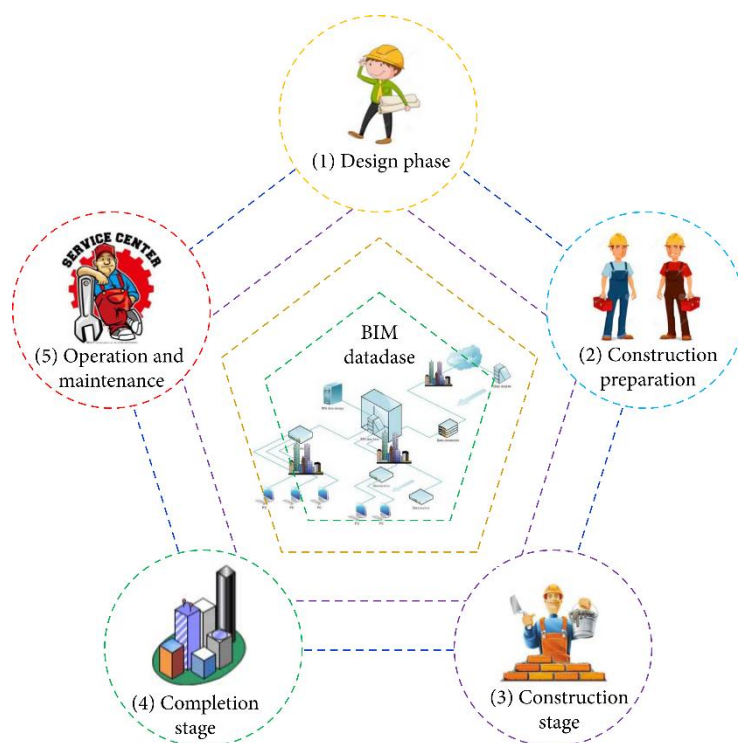


Рисунок 2. Источник данных BIM

Новый уровень управления проектами. Новый качественный уровень управления проектами в строительной отрасли сегодня связан с форматом международного сотрудничества, который достиг наивысшей степени развития. Использование 3D и 4D моделирования существенно упрощает и ускоряет различные рабочие процессы. Технология BIM позволяет специалистам обмениваться значительными объемами данных в любое время суток и в любой точке мира, выводя управленческую и проектную деятельность на новый уровень. В настоящее время использование офисов с открытой планировкой уже не так актуально и популярно. Эффективная командная работа становится возможной в онлайн-формате, когда специалисты находятся в разных географических точках, и это не отрицательно влияет на результаты. Кроме того, технология Big data представляет собой прогрессивный мониторинг проекта, способствующий повышению качества планирования, прогнозирования и оптимизации бюджета. Например, американская компания «Nick Savko & Sons», специализирующаяся на земляных работах, автоматизировала свои производственные процессы с помощью глобальных локаторов, которые позволили контролировать работу удаленно. Эти локаторы собирали информацию о производственных процессах, ошибках и простоях, что позволило улучшить работу организации и внедрить инновационные технологии. Пример проекта белорусских специалистов с использованием BIM технологий представлен на рисунке 3.



Рисунок 3. Одна из первых работ «Белпромпроекта», выполненных в BIM

Заключение. Применение самых передовых информационных технологий в любой отрасли обычно приносит положительные результаты. Однако использование нескольких информационных технологий одновременно может привести к существенному синергетическому эффекту. Строительная отрасль, из-за своего масштаба, предоставляет обширные возможности для применения таких технологий. В то же время, масштабность отрасли приводит к образованию огромного объема данных при использовании информационных технологий. Эти объемы данных часто слишком велики для анализа человеческими силами, и даже с привлечением большого коллектива людей такая работа может занять значительное время.

Применение технологии обработки больших данных в строительстве позволяет не только обрабатывать большие объемы данных, получаемые при использовании BIM-технологий, диспетчеризации и эксплуатации, но и объединять и анализировать эти данные в целом. Это создает еще более значительный синергетический эффект, чем простое сочетание этих технологий. Также это позволяет значительно повысить точность прогнозирования и управления за счет минимизации или уменьшения влияния человеческого фактора.

Список литературы

- [1] Волынский, В.Э. Большие данные в градостроительстве. 2017. – 324 с.
[2] Фольк Р, Штенгель Й, Шультманн Ф. Моделирование информации о зданиях для существующих зданий – обзор литературы и будущие потребности. 2014. – 204 с.
[3] Миеттинен Р, Паавола С. За пределами утопии BIM: подходы к разработке и внедрению моделирования информации о здании. 2014. – 104 с.

Авторский вклад

Мигалевич Сергей Александрович – руководство исследованием потенциала современных информационных технологий в области строительства, проектирования и эксплуатации строительных объектов.

Марков Алексей Николаевич – рассмотрение возможности совместного использования различных информационных технологий, технологий больших данных и BIM-технологий для решения задач в строительстве.

Богрецов Герман Андреевич – выявление потенциала повышения эффективности определенных строительных процессов за счет применения комбинированных информационных технологий.

IMPLEMENTATION OF BIG DATA PROCESSING INTEGRATION IN CONSTRUCTION

S.A. Migalevich

*Master of Technical Sciences,
Head of the Center for
Informatization and Innovative
Developments*

A.N. Markov

*Master of Technical Sciences,
Deputy Head of the Center for
Informatization and Innovative
Developments*

H.A. Bahratsou

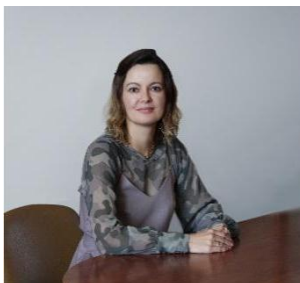
*Student of the Belarusian State
University of Informatics and
Radioelectronics*

Abstract. The aim of this study was to demonstrate the potential of modern information technologies in the field of construction, design, and operation of building structures. An additional objective was to explore the possibility of combining or jointly using various information technologies, such as big data technology and BIM technology, to address construction-related tasks. Furthermore, the study aimed to identify the potential for enhancing the efficiency of certain construction processes through the application of such combined information technologies.

Keywords: big data, construction processes, risk tracking, dispatching, operation automation.

УДК 338.22.021.4

ВОЗДЕЙСТВИЕ ИНФОРМАЦИОННЫХ И МЕДИАТЕХНОЛОГИЙ НА ТРАНСГРАНИЧНУЮ МОБИЛЬНОСТЬ КАПИТАЛОВ



О.И. Лаврова

Декан инженерно-экономического факультета БГУИР,
кандидат экономических наук, доцент
o.lavrova@bsuir.by

О.И. Лаврова

Окончила Белорусский государственный университет. Область научных интересов связана с исследованием проблем регулирования экономических процессов и рынков, в том числе с использованием идеологических институтов, а также особенностей применения информационных и медиатехнологий в решении актуальных макроэкономических задач.

Аннотация. В статье рассматриваются аспекты влияния информационных и медиатехнологий на различные экономические процессы с точки зрения безопасности и устойчивого развития национальной экономики. Перемещение способов информационного обмена в цифровой формат создает дополнительные угрозы в части оттока финансового и человеческого капиталов, что в ситуации идеологического разрыва может принять неконтролируемый масштаб и иметь негативные последствия в средне- и долгосрочной перспективе.

В условиях парадигмального сдвига от экономики консьюмеризма к экономике гуманизма, а также перехода «от моделей к методам», предусматривающего использование больших данных в целеполагании и принятии решений на макроуровне, требуется, чтобы меры экономической политики, основанные на алгоритмах искусственного интеллекта, соответствовали ценностным ориентирам своей страны, иначе потенциал манипулятивных инструментов, широко доступных в цифровой среде, может быть максимально эффективно реализован во благо развития других государств.

Ключевые слова: идеологический разрыв, отток ресурсов, информационная политика, гиг-экономика, большие данные.

Введение. Экономические процессы в современных условиях стали все активнее и интенсивнее разворачиваться в цифровой среде. В частности, благодаря открытости рынков и развитию глобальных сетевых коммуникаций в процессе обмена существенно возросло количество трансакций, которые, с одной стороны, стали более продуктивными, а с другой – информационно насыщенными. Цифровизация создала миф доступности информации как общественного блага, однако ее ценность как производственного ресурса обусловила возникновение противоположных тенденций в направлении закрытости и ограниченности доступа к данным, причем если предприятия преследуют преимущественно коммерческий интерес, то финансовые структуры и органы государственного управления руководствуются соображениями безопасности. В этой связи возникла парадоксальная ситуация, когда при наличии огромных массивов данных, именуемых в настоящее время «большими данными», проблема информационной асимметрии и дефицита знаний продолжает оставаться актуальной.

Искажение информационных потоков в экономическом и общественном пространстве. Международное экспертное сообщество обеспокоено возникновением

феномена «загрязнения» информационных потоков, сформированного вследствие попадания в коммуникационную среду непроверенных сведений либо заведомо ложной информации в целях манипулирования сознанием потребителей, а также поведением других социальных групп, в частности, участников биржевых торгов. В свое время советская культура полутонов и двойных смыслов была уничтожена тотальным распространением информации сомнительного содержания в результате политики гласности и свободы слова в перестроечный период. Кооперационные связи, поддерживаемые идеями коллективизма и солидарности, распались, как и социальный капитал, потенциал которого за 33 года так и не удалось полностью восстановить. Идеология «консюмеризма» доминирует в современной капиталистической системе независимо от способа координации экономических отношений – распределительного или рыночно-алгоритмического (то есть вне зависимости от типа X- или Y- экономики).

Индивидуализм, являющийся методологической основой неоклассики, характеризует человека рационального, реализующего свои материальные потребности и имеющего при этом склонность к оппортунизму. Однако идеологические институты, включающие ценностные ориентиры и морально-этические нормы поведения, приобретают все более существенное значение в определении потребительского выбора. Как отмечает академик Глазьев С.Ю., «искусственно создаваемые стереотипы играют в современном общественном сознании такую же роль, как обычаи и мифы в доиндустриальном обществе, – они влияют на пропорции обмена куда больше затрат труда и предельной полезности» [1, С. 544]. Запредельно высокое разнообразие товаров и услуг искажает объективное восприятие их ценности, в результате чего потребитель готов платить больше за приобретение брендовой продукции, наличие экологичной упаковки, псевдоинноваций, а в случае с белорусскими потребителями – импортных товаров «европейского» качества.

Ментальные ловушки политической коммуникации описаны Г.В. Пушкаревой, раскрывающей в работе [2] проблему попадания в колею группомыслия, из которой человеку трудно выбраться. Бойкот «западных» брендов, включая уход сети ресторанов Макдональдс из Беларуси и России, исследователи назвали культурой отмены, – настолько прочно в потребительском поведении укоренились поведенческие паттерны определенного формата, поэтому их разрушение создает угрозу экономической безопасности и социальной стабильности.

Обозначенная проблема усугубляется перемещением способов информационного обмена в цифровой формат, когда в процессе диалога упрощается аргументация, а анонимность собеседников обуславливает возможность открытой конфронтации и даже агрессии по спорным вопросам. Возникают предпосылки раскола в обществе, недоверия и нестабильности, что прежде всего негативно отражается на распределении финансовых ресурсов, доходов и инвестиционной привлекательности отраслей национальной экономики.

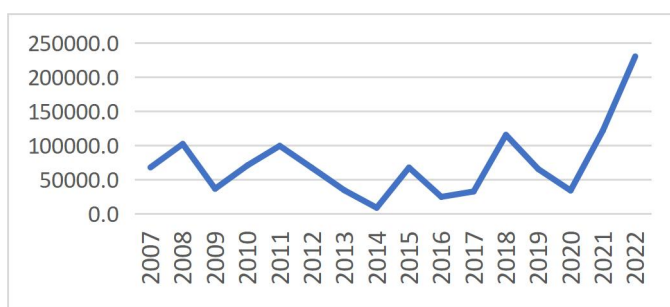
Отток капитала как системная проблема экономики в условиях идеологического разрыва. Поляризация общества как следствие обесценивания социального капитала (возникновение недоверия и разрушение горизонтальных связей) приводит к формированию экономики, функционирующей в условиях идеологического разрыва, когда система ценностей и взглядов не разделяется большинством населения страны (или большей частью ее элит). Е.Г. Егоровым идеология определяется как часть модели мира, отражающая социальную реальность и природу человека, и зачастую они не совпадают [3]. Исходя из данной логики, определим, что *идеологический разрыв* – это несоответствие действительности и истинных ценностных установок тем представлениям и убеждениям, которые транслируются в социуме и реализуются элитами в конкретном историческом контексте. На примере СССР понятно, что марксистская идеология не

вполне адекватно отражала реальную сущность и природу социально-экономических отношений, что в итоге привело к тотальной бюрократизации советской власти и пусть к неясному, но все-таки расслоению общества, о чем свидетельствуют предсмертные записки академика Е.С. Варги, написанные им в 1964 году [4].

Идеологический разрыв может возникнуть непреднамеренно по причине наличия добросовестных заблуждений и логических ошибок, а также сознательно в целях манипулирования общественным сознанием [3]. Экономическими последствиями такого разрыва будут угрозы нестабильности и безопасности в виде оттока ресурсов, как материальных (финансовых, трудовых), так и нематериальных (компетенций, знаний, технологий).

В частности, с 2021 года в связи с нарастанием геополитических противоречий и раскола в международном сообществе по российско-украинскому вопросу на основе данных баланса капитального счета можно наблюдать существенный отток капитала из страны (см. рисунок 1). Данные консолидированного счета операций с капиталом свидетельствуют о том, что Россия на регулярной основе поддерживает экономики других стран. Но в 2022 году объем чистого кредитования «остального мира» достиг рекордного за несколько последних десятилетий уровня – 230,4 млрд. долл. США, при этом доля валового накопления основного капитала в сбережениях упала до 65 %, что неудивительно в сложившейся ситуации.

Проблема оттока капитала является сверхактуальной для России, но если раньше его движение осуществлялось преимущественно в форме прямых инвестиций и соответствовало мировым трендам, то в условиях санкционного режима резко возросло наращивание объемов наличной валюты и депозитов, ссуд, займов и прочей задолженности (с 8,9 млрд. долл. США в 2019 году до 68 млрд. долл. США в 2022 году) [5]. Как отмечает А.С. Булатов, «усиление геополитических рисков становится важным, если не главным тормозом на пути потоков международного движения капитала» [5, С. 67]. Нарастание геоэкономической фрагментации в мире из соображений стратегии безопасности – современный тренд, который будет иметь долгосрочное будущее.



Источник: разработка автора на основе данных Росстата

Рисунок 1. Чистое кредитование «остального мира» России в 2007 – 2022 гг., млн. долл. США.

В Беларуси отток капитала находит отражение в других процессах, связанных прежде всего со слабой валютой и склонностью к долларизации экономики, а также с культурой потребления, характеризуемой паттерном «импортное лучше», который хотя и был сформирован исторически, но не без участия рекламы и новостей в СМИ, включая современные интернет-технологии. Рассмотрим кейс касательно внешнеторговых потоков Республики Беларусь. Потребительский импорт составляет четвертую часть всего странового импорта (9,3 млрд. долл. США) [6]. При этом не весь его объем отражается в платежном балансе страны, а только тот, который учитывается в декларациях и других таможенных документах. «Неофициальный» импорт оценивается Национальным

статистическим комитетом с применением балансового метода по результатам выборочных обследований домохозяйств. Если посмотреть на различия в статистических данных, публикуемых Национальным статистическим комитетом и Национальным банком, то можно заметить, что косвенная поддержка белорусским населением зарубежных экономик за последнее пятилетие составляла ежегодно порядка 2 – 3 млрд. долл. США, что сопоставимо с объемом всего белорусского экспорта компьютерных услуг в досанкционный период (см. таблицу 1). Проблемы геополитического характера (очереди на границе, сложности в получении визы и пр.) существенно повлияли на сокращение потребительского импорта, уменьшив вывоз валюты из страны почти вдвое до 1,2 млрд. долл. США в 2022 году. К сожалению, идеологические институты в этом процессе не были задействованы, поэтому в ментальном плане потребуется провести серьезную работу с населением и, в частности, с молодежью.

Таблица 1. Импорт товаров в 2018 – 2022 гг. по данным Национального банка и Национального статистического комитета, млн. долл. США

Показатель	2018 г.	2018 г.	2020г.	2021 г.	2022 г.
1 Импорт (данные Нацбанка)	35 932,6	36 525,7	30 388,6	39 789,1	37 259
2 Импорт (данные Белстата)	38 441,1	39 476,7	32 767,4	41 784,8	38 468
3 Разница (3=2-1)	2508,5	2951	2378,8	1995,7	1209
4 Доля «неофициального» импорта в страновом импорте, % (4=3:1)	6,5	7,5	7,3	5,0	3,2

Источник. Разработка автора на основе [6] и данных Белстата

Гиг-экономика, мобильность человеческого капитала и проблема «утечки умов». Гиг-экономика (или фриланс, частичная занятость) как новая форма экономических отношений стремительно развивается в современных условиях цифровизации и автоматизации, а также периодически сменяющих друг друга кризисах перепроизводства и структурных диспропорций. Прогнозируется, что в ближайшей перспективе гиг-занятых будет порядка 40 % всего занятого населения. Это явление провоцирует усиление тенденций мобильности человеческого капитала и создает угрозу деградации социального капитала, поскольку частично занятые работники не вовлечены в корпоративную культуру и их связи с другими людьми очень неустойчивы.

Использование информационных и медиатехнологий способно ослабить негативные эффекты гиг-занятости, поддерживая социальные связи в виртуальном пространстве. Фрилансер сегодня не ограничен рамками не только организации, но и территории (города, страны), поэтому государства, уступающие в своем социально-экономическом развитии лидерам, как правило из западного блока, без соответствующих мер государственной политики будут вынуждены сталкиваться с оттоком трудовых ресурсов на постоянной основе.

Самой серьезной угрозой является отток высококвалифицированных кадров из высокотехнологичных сфер экономики. В частности, по данным *NCSES* (США) за 2019 год удельный вес *STEM*-специалистов иностранного происхождения, работающих на постоянной основе в США и имеющих степень бакалавра и выше, составил почти 23 %, при этом большая часть из них имеет степень магистра (38 %) и даже доктора наук (45 %) [7]. Доля иностранцев с докторской степенью, занятых в академической сфере, также высокая – 33 % всех *STEM*-профессоров, что неудивительно, поскольку США по-прежнему остается основным реципиентом научных кадров со всего мира. Большинство из них являются уроженцами Китая (24 %) и Индии (15 %), эмигрантов из России гораздо

меньше – порядка 2 %. В последние годы политика Китая по вопросу «утечки умов» достаточно агрессивна, так, по данным американского агентства биржевых новостей Barron's в 2021 году обратно на родину вернулось почти 1400 китайских исследователей, работавших в STEM-лабораториях США [8]. Такая же ситуация с китайскими студентами, их приток на обучение в вузы США сократился почти на 50 % по сравнению с допандемийным уровнем.

Проблема оттока трудовых ресурсов существенна и для Республики Беларусь. Основываясь на результатах исследования Международной организации по миграции, ведущей профессиональной категорией работников, выезжающих за границу при содействии фирм-лицензиатов, являются лица, имеющие рабочие специальности [9]. Однако большая часть белорусов выезжает в поисках работы самостоятельно. Их динамика положительна в направлении как Российской Федерации, так и западных соседей, лидером среди которых является Польша. В 2021 году в эту страну в целях трудоустройства выехало 135,7 тыс. человек.

Безусловно, в процессах трансграничного движения капиталов (как финансового, так и человеческого) играют ключевую роль объективные экономические факторы: инвестиционный климат, разница в доходах, масштабы рынков и т.д. Однако в последнее время исследователи стали интересоваться иррациональными, а также экзистенциальными причинами поведения экономических агентов, изучением каузальных связей на основе экспериментального инструментария и больших данных.

Большие данные и информационная политика в интересах белорусской экономики. В настоящее время говорят о парадигмальном сдвиге в экономической науке «от моделей к методам» [10], т.е. к экономике данных, когда при помощи интеллектуальных систем в режиме реального времени прорабатываются эффективные регулятивные меры, повышается обоснованность целей и механизмов их достижения при оптимальном расходовании имеющихся ресурсов. Однако возможности искусственного интеллекта должны дополняться человеческой составляющей, агрегирующей качественную информацию, которая учитывает ценностные ориентиры и морально-этические установки общества, разделяемые элитой и транслируемые большинству.

Информация и информационная политика используется как инструмент воздействия на эмоции и инстинкты людей, их уязвимые места, манипулируя сознанием и взывая, например, к демонстративному потреблению, к поискам лучшего места проживания и работы, к определенным направлениям вложения своих капиталов. В исследовании российских экономистов на основе семантического анализа новостных текстов в «западных» СМИ рассматривалось влияние тональности и эмоциональной окраски новостей о России на динамику и структуру ее внешней торговли, начиная с 2014 года [11]. Авторы показали, что негативный новостной фон коррелирует и даже является фактором уменьшения внешнеторгового оборота России со странами ЕС.

Несмотря на то, что ментальные и социокультурные факторы экономического поведения, представленные неформальными институтами (привычки, традиции, общественные нормы), считаются достаточно устойчивыми и не подвержены значительным изменениям в конкретном историческом моменте, смена идеалов и ценностей периодически происходит как естественным, эволюционным, так и революционным путем. В первом случае изменения обусловлены поколенческими и демографическими сдвигами, крупными технологическими инновациями (массовое производство автомобилей, айфонomania и пр.), во втором – вызваны социальными потрясениями и кризисами. В частности, пандемия коронавируса, обострившая проблемы неравенства в капиталистических странах, заставила пересмотреть идеалы свободного рынка, основанные на выгоде и конкурентной борьбе, в пользу морали, альтруизма и справедливости.

Белорусский менталитет также претерпел изменения. Если раньше, в публикациях 10-20-летней давности белорус единодушно представлялся трудолюбивым, доверчивым, совестливым и гостеприимным крестьянином с явно выраженным коллективным началом, то сегодня современный соотечественник адаптировался к рыночным условиям и развил свои предпринимательские способности, что подтверждается данными официальной статистики: в 2022 году почти треть занятых в экономике Беларуси трудились на предприятиях малого и среднего бизнеса, а также являлись индивидуальными предпринимателями. В последнее время белорусские предприниматели стали активнее специализироваться в производственных отраслях с выраженной творческой составляющей и высокой долей ручного труда (украшения, предметы декора, дизайн, одежда и сумки).

Цифровизация и технологизация средств и каналов коммуникации может существенно воздействовать на устойчивость менталитета, ускорить и активизировать процессы разрушения морально устаревших ценностей, стать фактором, обуславливающим движение общества по траектории идеологического разрыва. Самой мощной властью обладают сегодня социальные сети, способные управлять общественным сознанием при помощи алгоритмов и при отсутствии идеологии направить массы в выгодном их владельцам направлении.

Эволюцию роли и места человека в экономике можно представить следующим образом:

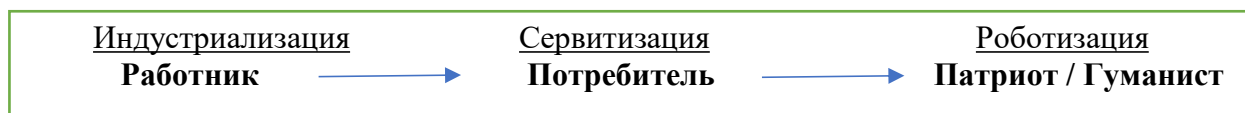


Рисунок 2. Стадии экономического развития и место человека в нем

В экономике потребления, развитие которой обусловлено процессами сервитизации и цифровизации, ценностные ориентиры эгоистичны и вызывают к инстинктам доминирования посредством демонстративного потребления, что всецело воспринимается населением независимо от «западного» или «восточного» типа общества, X- или Y-экономики. На современном этапе «западные» элиты нацелены на гуманизацию экономических отношений, вот только важно понять истинны ли мотивы и действительно ли человек реализует свое альтруистическое начало или это «тонкая» манипуляция общественным сознанием в целях достижения монопольного положения и получения сверхприбылей?

Вопрос дискуссионный, и если обратиться к институту собственности как инструменту власти, то на современном этапе она стала коллективной и перестала быть средством монопольного контроля над производственными активами (хотя ранее частная форма собственности считалась неизменным атрибутом капитализма). В этой связи неолиберальная идеология вместо идей свободного рынка и конкуренции транслирует по всему миру идеи и идеалы гуманизма и справедливости. Например, активно продвигается идея переселения людей из депрессивных районов в благополучные регионы, при этом не рассматриваются ни экономические, ни социокультурные последствия такого перераспределения человеческих ресурсов, а также позиции стран-доноров по этому вопросу. Какими средствами будет решена проблема неравенства – покажет время.

Заключение. «Загрязнение» информационных потоков, ангажированность СМИ, манипулятивные стратегии и коммуникационные ловушки – это все инструменты управления поведением потребителя в экономике консьюмеризма. Обосновано, что наиболее уязвимыми в части оттока финансового и человеческого капиталов являются страны, оказавшиеся в моменте идеологического разрыва, когда идеология либо

отсутствует, либо она устарела и больше не отражает ценностные ориентиры общества. Показано, что информационная и экономическая политика должны опираться на преимущества цифровизации и технологизации средств и каналов коммуникации, на работу алгоритмов в принятии стратегических и тактических мер (на принципах объективности, оптимальности, прозрачности), однако возможности искусственного интеллекта необходимо сочетать с концептуальными, качественно проработанными решениями, учитывающими социокультурные и ментальные установки, разделяемые элитой и большинством населения страны.

Список литературы

- [1] Глазьев С. Рывок в будущее. Россия в новых технологических и мирохозяйственных укладах. – М. : Книжный мир, 2018. – 768 с.
- [2] Пушкарева, В. Г. Политический менталитет и ментальные ловушки политической коммуникации / В. Г. Пушкарева // *Общественные науки и современность*. – 2023. – № 3. – С. 54 – 67.
- [3] Егоров, Д. Г. Модель человека как ключ к пониманию идеологии (о немарксистской альтернативе либерализму) / Д. Г. Егоров // *Общественные науки и современность*. – 2023. – № 3. – С. 23 – 38.
- [4] Варга, Е. «Скрыть через 25 лет»: заметки не для печати / Е. Варга // *Экономист*. – 2022. – № 12. С. 42 – 66.
- [5] Булатов, А. С. Новые тренды в движении капитала в мире и в России / А. С. Булатов // *Вопросы экономики*. – 2023. – № 9. – С. 65 – 83.
- [6] Платежный баланс Республики Беларусь // Национальный банк Республики Беларусь [Электронный ресурс]. – Режим доступа : <https://www.nbrb.by/statistics/balpay>. – Дата доступа : 20.01.2024.
- [7] Science and Engineering Indicators // National Center for Science and Engineering Statistics [Electronic resource]. – 2024. – Mode of access: <https://nces.nsf.gov/pubs/nsb20212/participation-of-demographic-groups-in-stem#foreign-born-workers-in-stem>. – Date of access : 25.01.2024.
- [8] Tang, Christopher S. The U.S. Is Losing Scientists to China. How to Stem the Flow [Electronic resource]. – 2024. – Mode of access: <https://www.barrons.com/articles/the-us-is-losing-scientists-to-china-competition-51664825686>. – Date of access : 22.01.2024.
- [9] Целостная оценка текущих стратегий экономического развития, трудовой миграции и денежных переводов в Республике Беларусь // Международная организация по миграции [Электронный ресурс]. – Режим доступа : <https://belarus.iom.int/sites/g/files/tmzbd11266/files/documents/Holistic-Assessment-Belarus-RUS.pdf>. – Дата доступа : 15.01.2024.
- [10] Клейнер, Г. Б. Экономика неплатежей: 30 лет спустя / Г. Б. Клейнер // *Вопросы экономики*. – 2023. – № 9. – С. 138 – 154.
- [11] Федорова, Е. А. Влияние новостного освещения России в СМИ на экспортно-импортную деятельность / Е. А. Федорова, С. О. Мусиенко, И. С. Демин, Ф. Ю. Федоров, Д. О. Афанасьев // *Вопросы экономики*. – 2019. – № 8. – С. 30 – 44.

THE IMPACT OF INFORMATION AND MEDIA TECHNOLOGIES ON CROSS-BORDER CAPITAL MOBILITY

O.I. Lavrova

*Dean of the Faculty of Engineering and Economics of BSUIR,
PhD of Economic Sciences, Associate Professor*

Abstract. The article examines the aspects of the influence of information and media technologies on various economic processes from the point of view of security and sustainable development of the national economy. The transfer of information exchange methods to a digital format creates additional threats in terms of the outflow of financial and human capital, which in a situation of an ideological gap can take on an uncontrolled scale and have negative consequences in the medium and long term.

In the context of a paradigm shift from the economy of consumerism to the economy of humanism, as well as the transition "from models to methods" involving the use of big data in goal-setting and decision-making at the macro level, it is required that economic policy measures based on artificial intelligence algorithms correspond to the value orientations of their country, otherwise the potential of manipulative tools widely available in the digital environment, it can be implemented as effectively as possible for the benefit of the development of other states.

Keywords: ideological gap, outflow of resources, information policy, gig economy, big data.

УДК 004.75

АВТОМАТИЗАЦИЯ ПРОЦЕССА ТЕСТИРОВАНИЯ И ОТЧЕТНОСТИ НА ПРОЕКТЕ WEB-INSYNC В ЗАО «АЛЬФА-БАНК»



У. М. Козак

Студент БГУИР, специалист по обеспечению качества ПО, ЗАО «Альфа-Банк»
ulyana.kozak.new@gmail.com



В. Г. Лунин

менеджер по обеспечению качества ПО ЗАО «Альфа-Банк»

У. М. Козак

Студентка Белорусского государственного университета информатики и радиоэлектроники. Работает в ЗАО «Альфа-Банк» в должности специалиста по обеспечению качества ПО.

В. Г. Лунин

Менеджер по обеспечению качества ПО, ЗАО «Альфа-Банк», координирующий направление тестирования на проектах физических лиц.

Аннотация. На крупных проектах покрытие автотестов растет с каждым днем и когда тестовых сценариев становится очень много, вести документацию и отчетность, а также следить за запусками автотестов из разных мест становится непросто, для решения данной проблемы на проекте *Web-Insync* ЗАО «Альфа-Банк» применяется специальная методология *TestOPS*, которая реализуется с использованием инструмента *Allure TestOPS*, что позволяет хранить всю документацию, запуски тестов и отчеты в одном месте и на базе этих данных получать графики, анализировать работу фреймворка и создавать динамическую документацию на базе автотестов.

Ключевые слова: *Allure TestOPS*, автоматизация тестирования, *Playwright*, динамическая документация

Введение. В быстро изменяющемся мире незамедлительная реакция на критические события может сэкономить десятки тысяч долларов или спасти чью-то жизнь. Огромное количество данных генерируется разными источниками ежесекундно. Анализируя эти данные, можно предсказывать проблемы и реагировать на них в реальном времени, избегая и предупреждая серьезные последствия. До недавнего времени еженедельные и ежемесячные аналитические отчеты полностью покрывали запросы бизнеса. В ЗАО «Альфа-Банк» специалисты по обеспечению качества занимаются не только ручным, но и автоматизированным тестированием. Зона ответственности у каждого специалиста зависит от платформы, которую он тестирует.

Web QA – занимается тестированием сайтов и браузерных приложений, используя разнообразное ПО, методологии и технологии. В зону ответственности входит проверка фронта (*UI*) и интеграции на уровне (*API*).

Mobile QA – занимается тестированием мобильных приложений, используя разнообразное ПО, методологии и технологии. В зону ответственности входит проверка фронта (*UI*) для *Android* и *iOS* платформ.

Зоны ответственности за тестирование на проектах *INSNC* показаны на рисунке 1.

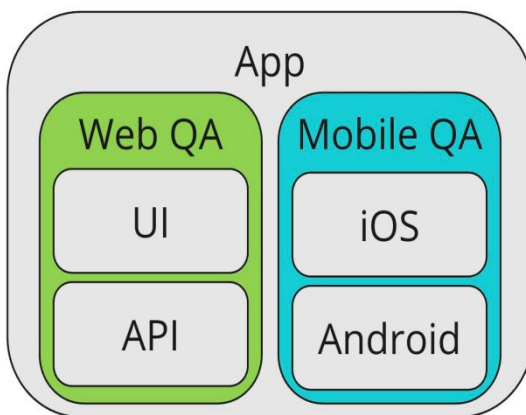


Рисунок 1. Зоны ответственности

Все тестирование проводится в соответствии с определенными этапами, начиная с тестирования бизнес-требований, далее тестирование макетов в *Figma*, тестирование *API* на ДЕВ, ТЕСТ стендах и написание автотестов на новый функционал, после тестирование интеграции бэка с фронтон и тестирование на пре-проду конечной реализации, оформление *UI* автотестов. Примерный жизненный цикл процесса тестирования показан на рисунке 2.

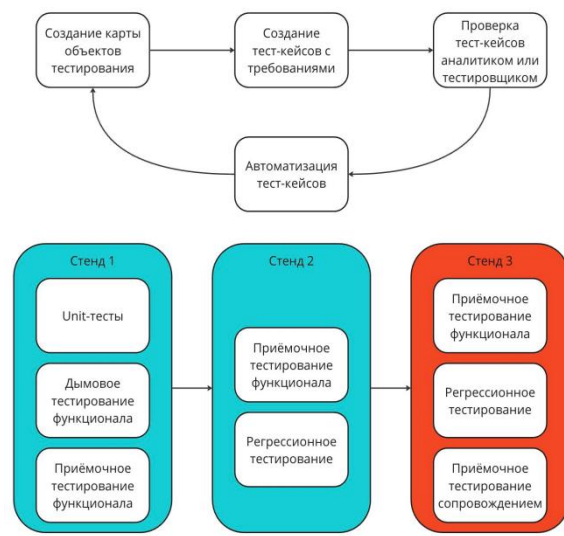


Рисунок 2. Примерный жизненный цикл процесса тестирования

В качестве основных инструментов для автоматизации тестирования используется следующий перечень приложений, показанный на рисунке 3. В центре всей схемы инструментария находится связующее звено, инструмент который позволяет сделать подход к тестированию унифицированным и организовать порядок в жизненном цикле разработки автотестов – *Allure TestOPS* [1].

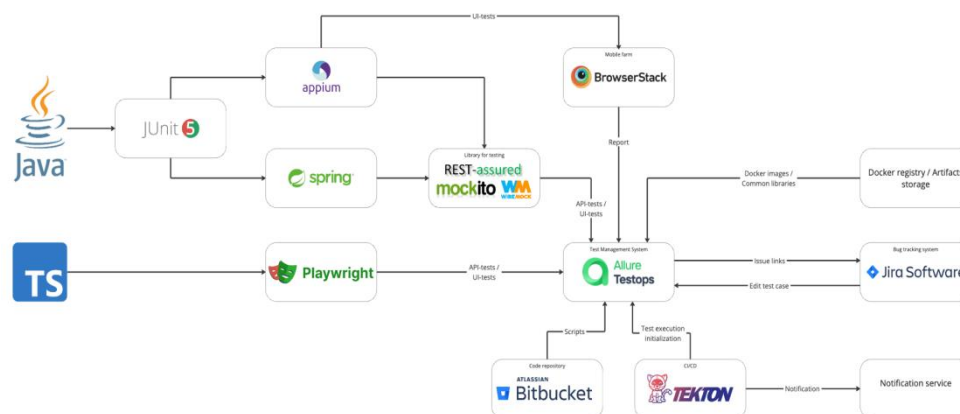


Рисунок 3. Инструментарий тестирования

В данной статье более подробно рассмотрим как использование в работе с большим количеством автоматизированных тестов и масштабными проектами, такими как *Web Insync* и *INSNC* мобильный использование *Allure TestOPS* помогает ускорить процессы работы не только с отчетностью но и с созданием документации.

Allure TestOps – это *Test Management System (TMS)*, которая изначально строилась с фокусом на автоматизацию и подход «тест-кейсы как код». Конечно же, это не случайно, т.к большинство автоматизаторов используют *Allure Report* и в коде автотестов уже используются механизмы разметки отчетов – эти же механизмы используются и для интеграции с *Allure TestOps* [2].

Allure TestOps стоит внедрять, если:

1 Написание и сопровождение подробной тестовой документации требует много ресурсов. В реальной жизни практически невозможно найти на это время. Например, сейчас в нашем тестовом проекте более 1000 *Playwright*-тестов и их число растет с каждым спринтом.

2 Тестовая документация плохо структурирована. *Allure TestOPS* помогает во первых отказаться от ручного написания тест-кейсов, а заменить этот процесс на автоматическую генерацию тест-кейсов на базе прогона, во вторых само приложение организовано так, что при добавлении в код авто-тестов специальных тегов можно указывать какая платформа, разработчик, протокол и т.п. соответствует данному кейсу, что в дальнейшем делает возможность поиска и фильтрации кейсов намного проще, а сами файлы – структурированными.

3 Отчеты хранятся в разных местах. После того, как тесты автоматизированы и готовы, мы запускаем их в *CI* (в Альфа Банке мы используем *TektonCI*). К каждой сборке прикреплен отчет *Allure*. Отчеты *Allure* это здорово, но по разным техническим причинам у нас много сборок, в которых прогоняются *Playwright*-тесты. Получалось так, что отчеты хранились не в одном месте, и проанализировать конкретный тест было сложно.

4 Проблемы со стабильностью тестов. Как и многие другие инженеры по автоматизации, мы мучились с нестабильными тестами. Но не все неудачные тесты являются нестабильными некоторые не проходят из-за ошибки в тестовой среде или просто требуют обновления. Выполнение этих тестов не только отнимало время у людей, просматривающих отчет, но и увеличивало время выполнения тестовой сборки. Было бы неплохо вообще не запускать эти тесты и исправлять их, когда придет время.

5 Не очевидно кто отвечает за тест. Благодаря специальным возможностям *Allure TestOPS* также можно в тегах указывать разработчика и тестировщика ответственного за данный тест-кейс, что помогает избежать недопониманий и конфликтов.

Чтобы понять, как нам удалось справиться со всеми проблемами, давайте посмотрим на код. В Альфа-Банке для написания автоматических тестов мы используем *JavaScript*, поэтому примеры кода также будут на нем.

Рассмотрим пример одного из наших простейших автоматических тестов:

Листинг 1: Пример UI автотеста с использованием Playwright

```
test.describe('Список истории операций', () => {
  test.beforeEach(async ({ authorizationSteps, pagesManager,
    responseUtils }) => {
    allure.owner('ukozak');
    allure.feature('История');
    allure.epic('История операций');
    allure.suite('Список истории операций');
    allure.labels(
      { name: 'Platform', value: 'Mobile' },
      { name: 'Platform', value: 'Desktop' },
      { name: 'layer', value: 'UI Tests' }
    );
    allure.tag('UI');
  });
  test('Отображение суммы зачисленных А-бонусов в истории
  @allure.id=23492', async ({
    pagesManager,
    historyVerificationSteps,
    textUtils
  }) => {
    await authorizationSteps.authorize(baseCredentialsModel);
    const [res] = await Promise.all([
      responseUtils.getResponseData(API_ROUTES.history.items),

    authorizationSteps.clickElement(pagesManager.headerPage.historyButton)
    ]);
    historyItems = res as HistoryItems;
    const { bonusSum, isBadgeRequired } = historyItems.items[29];

    //1 Проверить внешний вид суммы зачисленных бонусов bonusSum
    await historyVerificationSteps.checkElementText(
      pagesManager.historyPage.historyListBonusAmount,
      `+${textUtils.formatNumberThousands(bonusSum.amount)}${bonusSum.postfix}`
    );
    await historyVerificationSteps.checkCSS(
      pagesManager.historyPage.historyListBonusAmount,
      'color',
      COLORS_BONUS.positive
    );
    //2 Проверить отображение галочки на фронте для операции с параметром
    isBadgeRequired=true
    await historyVerificationSteps.checkElementPresence(
      pagesManager.historyPage.historyListBonusIcon,
      isBadgeRequired
    );
  });
});
```

Листинг 2: Пример API автотеста с использованием Playwright

```
import { test, *** } from '../..//dependencies';
test.describe('Детали операции', () => {
  let user: TestUserWithOperationModel, authHeaders: { [key: string]:
string };
  test.beforeEach(async ({ usersManager, authorizationSteps }) => {
    allure.owner('UKozak');
    allure.feature('Детали операции');
    allure.labels(
      { name: 'Platform', value: 'Web' },
      { name: 'Microservice', value: 'history-external-web' },
      { name: 'Developer', value: 'UKozak' },
      { name: 'layer', value: 'API Tests' },
    );
    allure.tag('RETENECC2C');
    allure.link('https://wiki.by/page', 'Протокол');
    user = (await usersManager.getUsersWithOperation(CurrencyType.BYN,
TypeOfOperation.TPD))[0];
    authHeaders = await authorizationSteps.firstLogin(user);
  });
  test('Отображение кнопки "Добавить в шаблон" в детальке операции
@allure.id=23239', async ({
  apiMethodsSteps,
  verificationSteps,
  historySteps
}) => {
    allure.tag('GET');
    allure.labels({ name: 'Endpoint', value: 'history/details' });
    const response = await apiMethodsSteps.getRequest(
      `${API_ROUTES.history.details}?id=${user.operationId}`,
      authHeaders
    );
    await verificationSteps.checkResponseStatus(response, StatusCode.OK);
    await verificationSteps.checkResponseSchema(response,
HISTORY_DETAILS_SCHEMA);
    await verificationSteps.checkResponseFieldValue(response, 'id',
user.operationId);
    await historySteps.checkButtonFromShowButtons(response, 'RECEIPT');
    await historySteps.checkButtonFromShowButtons(response,
'ADD_TO_TEMPLATE');
  });
});
```

Для того что бы на проекте можно было использовать *Allure TestOps* для отображения отчетов при использовании *Playwright*, необходимо предварительно установить *allure-playwright*.

1 Пример присвоения аннотаций ко всем тестам при помощи *beforeAll*

Листинг 3: Присвоения аннотаций *beforeAll*

```
test.beforeAll(async () => {
  allure.owner(ukozak);
  allure.feature('Шаблоны: создание, редактирование');
  allure.label({'name': 'UI', 'value': 'UI Tests'});
```

```
allure.link({ url: "****", name: "****" });
});
```

2 Название *suite*

По умолчанию репортер использует путь к тестовому файлу в качестве имени пакета. Если тесты используют *allure.suite()* и его значение должно использоваться в пользовательских полях *Allure TestOps*, установите опцию *suiteTitle: false*.

Листинг 4: Название *suite*

```
const config = {
  reporter: [['allure-playwright', {
    detail: true,
    outputFolder: 'my-allure-results',
    suiteTitle: false
  }]],
};
```

3 Пример использования *allure.label*

Аннотация *label* используется при работе кастомными полями. *Allure Labels* – это набор пар ключ-значение, которые используются для хранения метаданных в результатах *Allure*. Поддерживаемые типы атрибутов показаны в этой статье

Листинг 5: Пример использования *allure.label*

```
//Если одно значение:
allure.label({ 'name': 'Platform', 'value': 'Mobile' });
//Если два и более:
allure.label([
  { 'name': 'Platform', 'value': 'Mobile' },
  { 'name': 'layer', 'value': 'UI Tests' }
]);
```

Чтобы увидеть методы в отчете, необходимо использовать аннотацию *@Step*. Эта аннотация также может работать с параметрами, устанавливая фактическое значение переменной вместо заглушки.

Шаблон оформления тест-кейса в *Allure TestOps*, пример тестового сценарий представлен на рисунке 4, где цифрами обозначены пункты шаблона:

- название тест-кейса на русском языке и из него должно быть понятно, в чем суть тест-кейса;
- в поле *Link* указываем ссылки на спецификацию (Альфа-Знания) и структуру *API* (*Swagger*, *Google Docs* и/или др.);
- в поле *Description* записываем: для *API Tests* – Метод и *Endpoint* запроса, для *UI Tests* – Описание тест-кейса (необязательно);
- в поле *Precondition* записываем предусловия которые предшествуют данному запросу или другу важную информацию необходимую для выполнения данного запроса (Например: для *UI Tests* можно указать необходимые данные или шаги которые нужно выполнить заранее, чтобы выполнить тест-кейс (например создание предчека, выбор/создание необходимого пользователя и т.д.));
- в *Scenario* пишем проверки, которые необходимо провести. По необходимости можно писать в подшагах ожидаемый результат с ключевым словом [Ожидаю] либо [OP];

– в поле *Issues links* указываем ссылки на задачи в Альфа-Проекты связанные с данным запросом указываем *Test layer* указываем *API Tests* либо *UI Tests*.

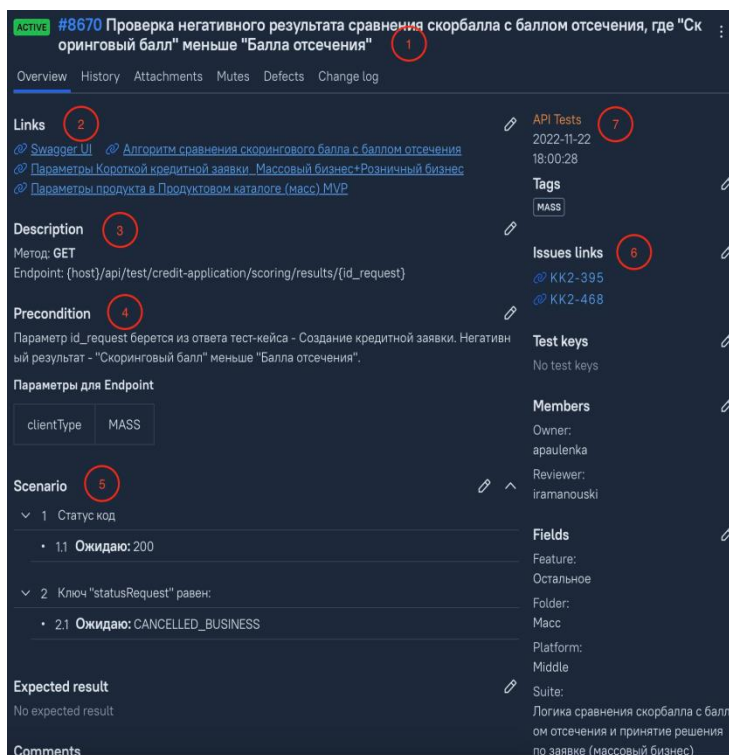


Рисунок 4. шаблон тест-кейса в *Allure TestOPS*

Концепции которых придерживаются на проекте *INSNC*.

1 Отчетность.

Отчет *Allure* отображает шаги и ошибки в тесте четко и понятно, что делает процесс тестирования более информативным и продуктивным. Все действия теста должны быть отражены и легко читаемы.

2 Структурированность.

Каждая папка проекта содержит свой логический слой данных или файлов, точно также каждый файл, следуя принципу единственной ответственности, должен содержать только связанный логикой набор функций и переменных.

3 *Page Object*.

Один из наиболее полезных и используемых архитектурных решений в автоматизации. Данный шаблон проектирования помогает инкапсулировать работу с отдельными элементами страницы, что позволяет уменьшить количество кода и его поддержку. Если, к примеру, дизайн одной из страниц изменен, то нам нужно будет переписать только соответствующий класс, описывающий эту страницу.

4 Уникальность проверок.

Идея заключается в том, что если функциональность проверена одним тестом, то нет смысла проверять ее же следующим тестом. Таким образом это помогает избавиться от лишнего кода, сократить время на потенциально дублирующие друг друга проверки и делает тесты более стабильными, поскольку выполняется меньше лишних действий [3].

5 Шаги.

Все действия направленные на создание каких-либо сущностей или на переход к какой-либо странице приложения выделяются в отдельную прослойку «*steps*», т.к. потенциально могут быть использованы многократно. Такой подход позволяет:

- избавиться от дублирования кода, в частности не требуется каждый раз писать *title* шага;
- отчет содержит подробную инструкцию о том, что конкретно выполнялось в рамках теста.

6 Предсказание ошибок.

В *Allure TestOps* можно оформить категории вероятных ошибок, что помогает реализовать концепцию *Error Guessing* («Предсказание ошибок»). Суть данного подхода заключается в том, чтобы при анализе и разборе упавших тестов ускорить процесс разбора причин возникновения упавших тестов: *Allure* разбивает кейсы запуска по определенным категориям, которые помечаются лейблами, на рисунке 5 приведены те категории которые на проекте *Web-Insync* используются для разбора упавших *UI* автотестов.

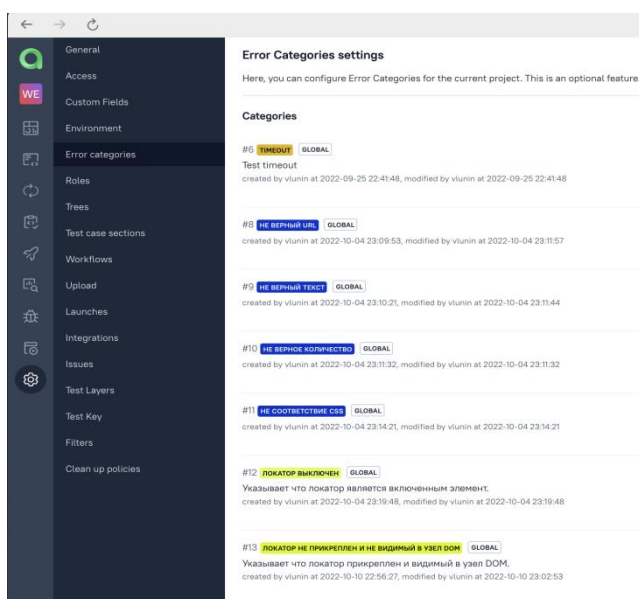


Рисунок 5. Категории ошибок на проекте *Web-Insync*

7 Деревья.

Allure TestOps представляет тест-кейсы с помощью древовидной структуры. Вы можете создавать различные деревья в своем проекте в зависимости от информации, представленной в пользовательских полях ваших тестов, и ваших потребностей. На проекте *Insync* разделение производится по *Feature*, *Platform*, *Suites*. На рисунке 6 приведен пример дерева фильтрации тестовых сценариев на проекте *Web-Insync*.

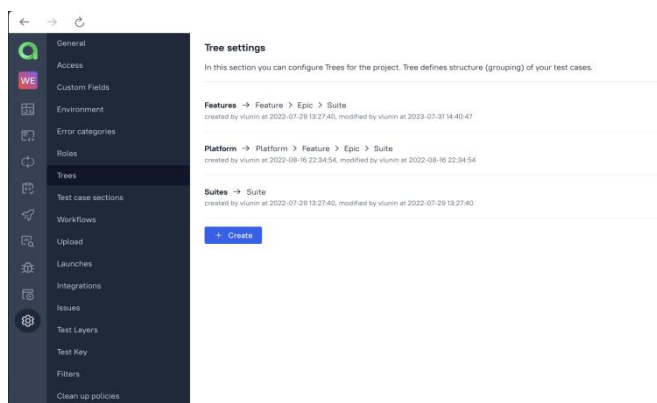


Рисунок 6. Деревья сценариев на проекте *Web-Insync*

8 Дешборды (Графики)

На проекте применяется 2 основных типа дашбордов - стандартные, которые по умолчанию присутствуют в разделе *Dashboard* в *Allure TestOPS*, и пользовательские, которые создаются по нуждам проекта в дополнение к стандартным [4].

Standard dashboard (стандартная панель) предоставляет мгновенное представление о статусе тестов, которые проводятся на проекте, и находится данная панель в разделе *Dashboards* проекта. Пример стандартной панели на проекте *Web-Insync* приведен на рисунке 7.



Рисунок 7. Стандартная панель *Web-Insync*, покрытие тестами

Как видно из рисунка на проекте достаточно большое количество тестовых сценариев, как мануальных, так и автоматизированный и данная панель помогает наглядно видеть результаты работы над тестированием на проекте.

С помощью стандартной панели можно увидеть следующие сведения:

- количество тестовых случаев и их распределение по состояниям (активные, на рассмотрении, устаревшие и т. д.);
- количество тестовых случаев и распределение между ручными и автоматизированными тестами;
- тренд автоматизации с информацией о количестве тестов за последние 14 дней;
- количество тестов в день и их статус (провален, пройден);
- продолжительность выполнения тестов в день;
- тренд Mutes, т.е. количество тестов, исключенных из статистики из-за постоянного состояния провала или состояния *flaky*.

Целью пользовательских панелей является предоставление дополнительных аналитических данных для лиц, принимающих решения. Конечный пользователь создает пользовательскую панель в разделе панелей проекта (кнопка + *New Dashboard* в правом верхнем углу). Дополнительная аналитическая информация предоставляется с помощью виджетов. Пользовательские панели на проекте *Web-Insync* приведены на рисунках 8-9.



Рисунок 8. Пользовательская панель *Web-Insync*, покрытие тестами

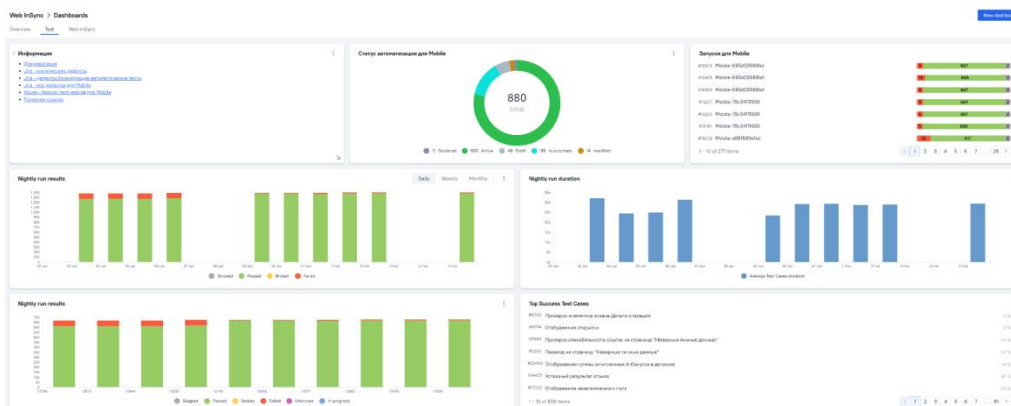


Рисунок 9. Пользовательская панель *Web-Insync*, общая информация

9 Динамическая документация

На проекте *Web-Insync* принятой практикой для создания тест-кейсов является генерация тестовой документации (тест-кейсов) на базе прогона автотестов. На рисунке 10 приведен пример тест-кейса сгенерированного на базе автотеста [5].

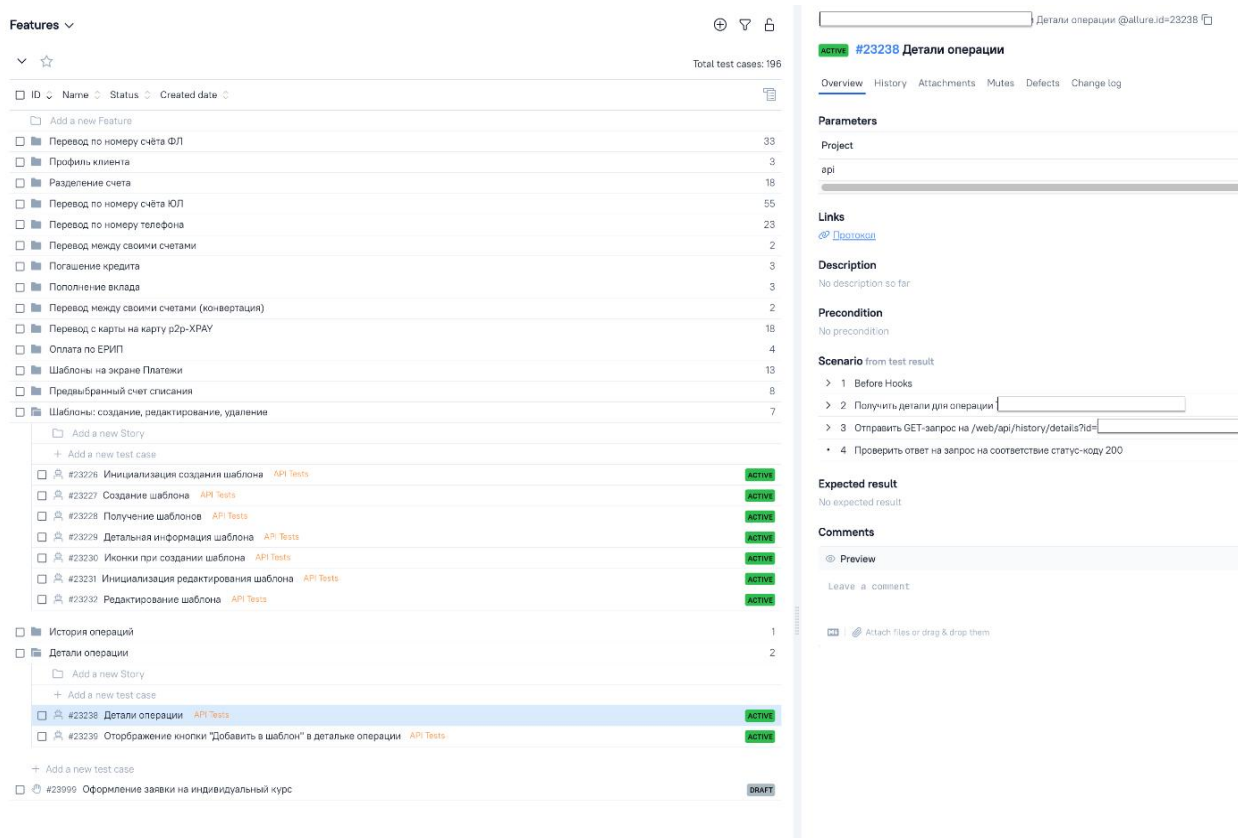


Рисунок 10. Пример тест-кейса, сгенерированного на базе кода автотеста

Как уже было описано ранее, весь код автотестов при написании тестировщик оборачивает в специальные блоки – *Steps*, которые *Allure* считывает и оформляет в понятный человеку набор проверок с ожидаемым результатом.

10 Интеграция *Allure* с *Grafana*

На крупных проектах необходимо управлять тестированием и принимать решения о релизах и понимать, есть ли в данный момент какие-то проблемы или нет. В ходе работы над данным процессом, возникает ряд проблем:

- разработчики пренебрегают изучением отчётов о тестировании;
- мало кто заглядывает в *Allure*, а если заглядывают, то не вполне понимают локаторы и сами сценарии;
- руководителям не хочется погружаться в детали – они хотят иметь картинку перед глазами;
- недостаточное покрытие и как его оценить;
- недостаточная скорость работы тестов.

Чтобы преодолеть эти трудности, нужно научиться измерять работу тестов. В этом помогает *Grafana* – инструмент, которым пользуются разработчики, девопсы, аналитики, продакт-менеджеры. *Grafana* отображает графики и результаты аналитики [6].

11 Интеграция *Allure* с *Jira*

Настроив интеграцию *Allure TestOps* с *Jira*, создается отдельная страница со всеми запусками (*Launches*) в *Allure TestOps*, пример данной страницы показан на рисунке 11. Страница *Launches* показывает эту информацию в структурированном виде, позволяет отфильтровать ее и получить нужный срез данных (запущенные тесты, процент успешно выполненных и упавших тестов и другую схожую информацию) по конкретной ветке.

Если провалиться в конкретный *Launch*, можно посмотреть на структуру проекта: какие есть компоненты, какие фичи, сколько было перезапусков, какие именно тесты перезапускались, информацию о дереве тест-кейсов, график, который показывает сколько времени бежал самый долгий тест, график с длительностью тестов, – и в конце концов экспортировать этот прогон в *JIRA*-тикет. Отдельно хочется упомянуть про политики для *Live documentation*, которая в автоматическом режиме обновляет дерево тест-кейсов по правилам и триггерам, например, только из ветки *master*.

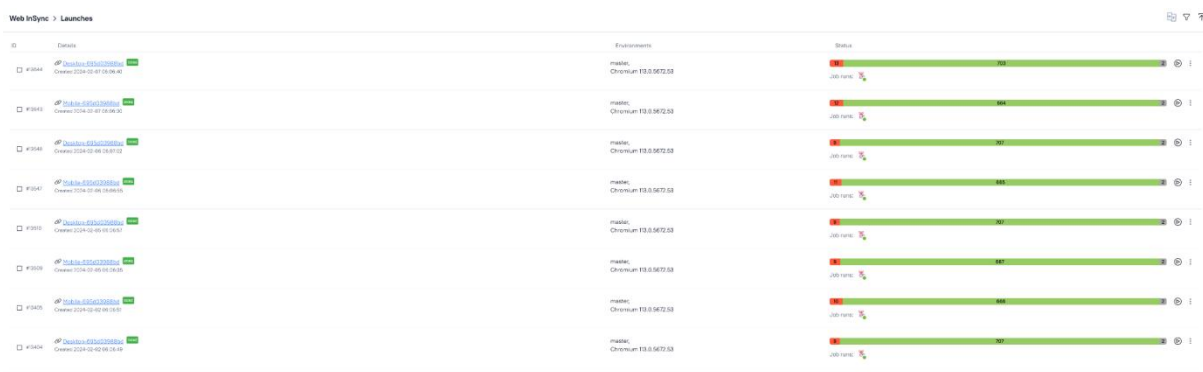


Рисунок 11. Страница запусков *Web-Insync*

Работает это так: у нас есть автотест, в котором изменили пару шагов и поправили аннотации. Далее этот тест запускается и выполняется с набором тестов на *master-branch*, после чего можно либо вручную закрыть *launch*, либо дождаться, когда *launch* будет закрыт по *cron*’у (расписание запусков), который настраивается в интерфейсе *Allure*. При закрытии запуска старый автотест обновляется - получаем актуальный тест-кейс. Это позволяет забыть про актуализацию тестовой документации в интерфейсе [7].

Заключение. Использование в тестировании на больших проектах, с множеством тестовых сценариев автоматизации делает работу быстрее, а анализ данных и метрик намного проще. В ЗАО «Альфа-Банк» тестирование проводится в соответствии с методологией *TestOPS*, где тестировщики выполняют работу как с по мануальному так и

автоматизированному тестированию функционала, а также используют в работе *Allure TestOPS* для автоматизации процесса отчетности и создания тестовой документации. Обработка и анализ тестовых метрик, таких как покрытие авто-тестами и результаты упавших тестов структурируются в *Allure Dashboards*, что позволяет бизнесу более наглядно анализировать результаты тестирования, а тестировщикам оценивать скоуп своей работы и планировать расширение фреймворка на базе имеющихся сведений. Также данный способ организации тестовых сценариев делает процесс регрессионного тестирования в конце спринта проще, что ускоряет работу, а соответственно экономит деньги.

Список литературы

- [1] Инструменты автоматизации тестирования [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/768154/>
- [2] Как мы внедряли Allure TestOps в стриминговом сервисе [Электронный ресурс]. – Режим доступа: https://habr.com/ru/companies/ru_mts/articles/689330/
- [3] Паттерны автоматизации и архитектура автотестов [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/777262/>
- [4] Dashboards Allure TestOPS [Электронный ресурс]. – Режим доступа: <https://docs.qameta.io/allure-testops/briefly/dashboards/>
- [5] QA Meta blog [Электронный ресурс]. – Режим доступа: <https://qameta.io/blog/>
- [6] Grafana и автотесты: учимся измерять работу тестов [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/companies/ozontech/articles/657933/>
- [7] Allure. В поисках почти идеальной TMS [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/571476/>

Авторский вклад

Авторы внесли равноценный вклад в написании статьи

AUTOMATION OF THE TESTING AND REPORTING PROCESS ON THE WEB-INSYNC PROJECT AT ALFA-BANK

U. M. Kozak

*BSUIR student, FullStack QA at AlfaBank,
Republic of Belarus*

V. G. Lunun

*Software Quality
Assurance Manager
at AlfaBank,
Republic of Belarus*

Annotation. On large projects the coverage of autotests is growing day by day and when there are a lot of test scenarios, it becomes difficult to keep documentation and reports, as well as to keep track of autotest runs from different places. To solve this problem on the Web-Insync project of CJSC «Alfa-Bank» a special TestOPS methodology is used, which is implemented using the Allure TestOPS tool, which allows to store all documentation, test runs and reports in one place and on the basis of these data to get graphs, analyze the work of the framework and create a dynamic analysis.

Keywords: Allure TestOPS, test automation, Playwright, dynamic documentation

УДК 004.021:004.75

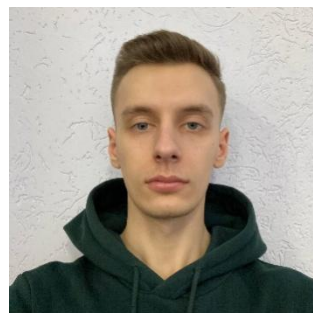
ИНТЕГРАЦИЯ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ В АВТОМОБИЛЬНОЙ ПРОМЫШЛЕННОСТИ



С.А. Мигалевич
Магистр технических наук,
начальник центра
информатизации и
инновационных разработок
migalevich@bsuir.by



А.Н. Марков
Магистр технических
наук, заместитель
начальника центра
информатизации и
инновационных
разработок
a.n.markov@bsuir.by



Д.Г. Ершов
Студент Белорусского
государственного
университета
информатики и
радиоэлектроники
d.ershov@bsuir.by

С.А. Мигалевич

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов: вибродиагностика, разработка метода вейвлет-анализа изделий машиностроения.

А.Н. Марков

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов: вычислительные системы, облачные вычисления (CLOUD COMPUTING), распределенные вычислительные системы, балансировка нагрузки вычислительных систем (load balancing).

Д.Г. Ершов

Студент Белорусского государственного университета информатики и радиоэлектроники факультета компьютерного проектирования.

Аннотация. По мере того как технология Интернета вещей (*IoT*) становится все более важным трендом для будущего транспорта, разработка крупных систем *IoT* становится критической задачей, направленной на обработку больших данных, загружаемых автопарками, и предоставление сервисов на основе данных. Данные *IoT*, особенно статусы транспортных средств с высокой частотой (например, местоположение, параметры двигателя), характеризуются большим объемом с низкой плотностью значений и низким качеством данных. Такие характеристики создают проблемы для разработки приложений в реальном времени на основе таких данных. В этой статье мы рассматриваем проблемы проектирования масштабной системы *IoT*, описывая *CarStream*, промышленную систему обработки больших данных для услуг аренды автомобилей с водителем. Подключенная к более чем 30 000 автомобилям, *CarStream* собирает и обрабатывает несколько типов данных о вождении, включая статус автомобиля, активность водителя и информацию о поездках пассажиров. На основе собранных данных предоставляются несколько сервисов. *CarStream* была задействована и поддерживается в промышленном использовании в течение трех лет, собрав более 40 терабайт данных о вождении. В данной статье рассмотрен опыт проектирования *CarStream* на основе потоков данных о вождении крупномасштабного характера, а также уроками, извлеченными из процесса решения проблем при проектировании и поддержке *CarStream*.

Ключевые слова: большие данные, автомобильная промышленность, отслеживание рисков, управление знаниями, автоматизация эксплуатации.

Введение. В последние годы технология Интернета вещей (*IoT*) стала важной областью исследований и применения. В качестве одного из основных направлений *IoT* Интернет вещей для транспорта (*IoV*) привлек большое внимание исследователей и промышленности. Недавно облачные технологии *IoV* получили выгоду от быстрого развития мобильных сетей и технологий обработки больших данных. В отличие от традиционных технологий сетей транспортных средств, которые фокусируются на взаимодействии транспортных средств (*V2V*) и сетях транспортных средств, в типичном облачном сценарии *IoV* транспортные средства подключены к облачному центру данных и передают статусы транспортных средств в центр через беспроводные коммуникации. Облако собирает и анализирует переданные данные и отправляет обратно на транспортные средства дополнительную информацию со значением. Аналогично другим приложениям *IoT*, данные о транспортных средствах обычно организованы в виде потока. Хотя каждое транспортное средство загружает небольшой поток данных, в облаке объединяется большой поток из-за как высокой частоты, так и большого масштаба флота. Поэтому основным требованием в этом облачном сценарии *IoV* является обработка потока больших данных о транспортных средствах в своевременном режиме (рисунок 1).

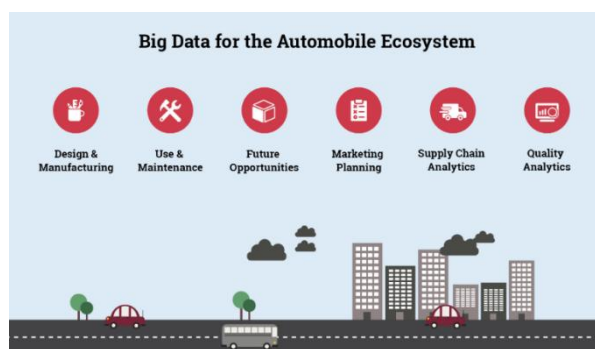


Рисунок 1. Большие данные в автомобильной промышленности

Существует несколько решений для проектирования системы *IoV*. Например, традиционная архитектура *OLTP* (*Online Transaction Processing*) использует зрелую и стабильную базу данных в качестве центра для развертывания сервисов. В такой архитектуре система собирает данные, загруженные транспортными средствами, и сохраняет их в базе данных, предоставляя дополнительные сервисы на основе аналитической способности подсистемы базы данных. Это решение просто, зрело и может быть очень надежным. Однако оно не масштабируется хорошо с ростом размера флота, потому что база данных легко может стать узким местом всей системы. Архитектура *OLAP* (*Online Analytical Processing*) часто используется для работы с крупномасштабными данными, в которой подсистема обработки данных, а не подсистема хранения данных (например, подсистема базы данных), предоставляет основную аналитическую способность. В такой системе нагрузка на вычисления в основном ложится на подсистему обработки данных, которая часто выполняет сложные запросы. Проектирование подсистемы обработки данных в стиле *OLAP* для сценариев *IoV* часто требует интеграции нескольких платформ, таких как *Storm* и *Spark Streaming* [1].

В данной статье рассмотрены задачи проектирования масштабной и высокопроизводительной системы аналитики больших данных для *IoV*, описывая *CarStream* – промышленную систему обработки данных для услуг аренды автомобилей с водителем. *CarStream* соединил более 30 000 автомобилей в более чем 60 городах и имеет несколько источников данных, включая данные о состоянии транспортного средства (такие как скорость, траектории, обороты двигателя в минуту (об/мин), оставшееся

количество бензина), активности водителя (например, начало обслуживания, прием пассажира) и заказы пассажиров. *CarStream* предоставляет несколько услуг в реальном времени на основе этих данных.

В частности, решается проблема масштабируемости, оснащая *CarStream* способностью к распределенной обработке и хранению данных. Затем используется подсистема обработки потоков для предварительной обработки данных на лету, чтобы решить проблемы низкого качества данных и низкой плотности значения. Данное решение достигает более высокой производительности за счет большего объема хранения. Также ускоряется работа *CarStream* дальше, интегрируя подсистему кэширования в памяти. Наконец, разрабатывается трехуровневую систему мониторинга для *CarStream* для обеспечения высокой надежности и управляемости. Система мониторинга обеспечивает всестороннее представление о том, как работает вся система, включая кластерный уровень, уровень вычислительной платформы и уровень приложения. Эта подсистема помогает в реальном времени определять состояние здоровья *CarStream* и также предоставляет ценную информацию о системе разработчикам, непрерывно улучшающим надежность и производительность системы. *CarStream* был развернут и поддерживается на протяжении трех лет в промышленном использовании, в процессе которого продолжается совершенствование и обновление системы, а также развиваются новые приложения.

Сбор больших данных. При обсуждении сбора данных для данного анализа вглубь больших данных вовлекаются различные методы сбора информации. Например, в автомобильной промышленности используются разнообразные инструменты, такие как GPS, датчики и камеры, установленные в транспортных средствах. Все эти компоненты генерируют структурированные, неструктурированные и полуструктурированные данные (всего 5–10 процентов всех данных). Для обработки и анализа данных, полученных из этих источников. Полученные данные обрабатываются и объединяются для предоставления различных услуг, таких как прогнозирование движения автомобилей для компаний, предоставляющих телематические услуги, страховых компаний и агентств по прокату автомобилей. Эта информация используется для понимания спроса и предложения на продукты и услуги компаний, что позволяет им предоставлять клиентам более индивидуализированный и персонализированный опыт. Использование таких данных создает новую экосистему, улучшающую взаимодействие с пользователями.

Нахождение универсального инструмента, подходящего для всех сценариев обработки данных, представляет собой значительную сложность. Многие предприятия постоянно стремятся создать эффективное хранилище и обработчик больших данных, который также был бы удобен для разработчиков. Однако одним из ключевых подходов, обеспечивающих более управляемую обработку, является разделение данных на более мелкие фрагменты и их параллельная обработка. Применение метода "разделяй и властвуй", широко используемого в информатике, эффективно и для работы с большими данными.

Предыстория *CarStream*. Концепция *IoV* (Интернет вещей в автомобиле) вытекает из *IoT* (Интернет вещей). В традиционном понимании, *IoV* описывает сеть транспортных средств, соединенных через *RFID* (радиочастотную идентификацию). Технология *V2V* (*Vehicle to Vehicle*) дополнительно обменивается данными о трафике и состоянии автомобилей, обеспечивая решения на местном уровне для безопасности и эффективности транспортировки.

С развитием мобильного интернета, традиционное *IoV* перешло к развертыванию сети широкой зоны, объединяющей внутренние и межавтомобильные сети. Датчики автомобилей подключены к электронному блоку управления через *CAN-BUS*, а облачная сеть автомобилей строится с подключением к серверам через мобильную сеть [2].

Система *IoV* позволяет собирать, анализировать и хранить данные автомобилей на серверах, которые также отправляют обратно автомобилям полезную информацию.

IoV стал одним из самых популярных приложений в *IoT*, однако крупные развертывания *IoV* редки из-за высоких затрат на обслуживание и развертывание. Приватные автомобильные службы, такие как *Uber* и *Lyft*, успешные примеры таких систем. В этой статье обсуждаются проблемы проектирования облачной системы *IoV* под названием *CarStream*, которая основана на китайской компании *UCAR*. Компания предоставляет аналогичные услуги, но с флотом, принадлежащим самой компании, и водителями, являющимися ее сотрудниками. Для сбора данных используется *OBD* (*Onboard Diagnostic*), который передает информацию на сервер через беспроводной модуль (рисунок 2).

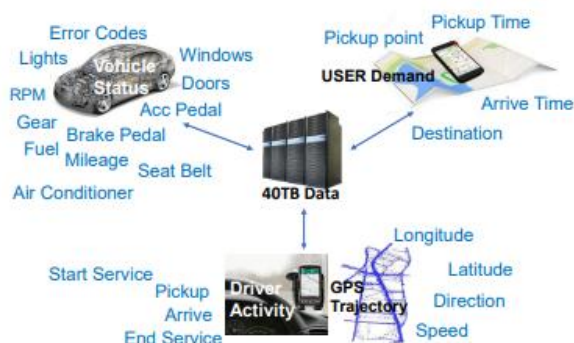


Рисунок 2. Устройство системы *CarStream*

CarStream обрабатывает данные о транспортных средствах и бизнесе в качестве сервиса на стороне сервера. На данный момент большая часть парка, более 30 000 транспортных средств, подключена к *CarStream*. Эти транспортные средства распределены по 60 различным городам Китая. Каждое транспортное средство загружает пакет данных на сервер каждые 1–3 секунды во время движения. Таким образом, серверная система должна обрабатывать почти 1 миллиард экземпляров данных в день. Объем парка расширился с 1500 до 30 000 за несколько месяцев; это расширение бизнеса создает запрос на горизонтальное масштабирование базовой платформы. В течение 3 лет *CarStream* собрало и обработало 40 ТБ данных о транспортных средствах.

Подсистема управления потоковыми данными. *CarStream* должен управлять терабайтами данных о поездках и предоставлять сервисы на основе анализа данных. Поток формируется большим количеством экземпляров данных, каждый из которых занимает сотни байт. *CarStream* ежедневно собирает почти миллиард экземпляров данных. Плотность полезной информации в таких данных очень низкая, так как из-за высокой частоты выборки можно извлечь лишь немного полезной информации. Кроме того, ценность данных быстро падает с течением времени, потому что большинство приложений, особенно те, что требуют временной чувствительности, в *IoV* более заинтересованы в новых данных.

Решение проблемы хранения данных в *IoV* с двух точек зрения. Первый аспект – это разделение данных. Выделяется приложения с высокими требованиями к реальному времени, чтобы далее выделить соответствующие «горячие» данные (например, текущий статус транспортного средства) отделяются «горячие» данные от «холодных» и помещаются в различные хранилища с разной производительностью. Второй аспект – это предварительная обработка. Сканирование больших наборов данных для базы данных затруднительно, но добавление данных (запись) может быть быстрым. Поэтому извлекается небольшой набор данных с более высокой плотностью ценности с помощью

потоковой обработки. Затем мы отделяем эту часть данных от сырых данных. Таким образом, подсистема хранения может обеспечить высокую пропускную способность для записи данных и высокую производительность для запросов, основанных на обработанных данных. Кроме того, мы помещаем некоторые «горячие» данные в кэш в памяти для обеспечения доступа к данным в реальном времени [3].

Для создания подсистемы управления данными мы использовали три различных платформы хранения в *CarStream*, включая кэширование в памяти, реляционные базы данных и хранилища больших данных *NoSQL*. Существует множество вариантов для каждого типа хранилища. Многие *NoSQL* базы данных, такие как *Cassandra*, *HBase* и *MongoDB*, могут быть использованы. На практике мы используем *HBase* в качестве архивного хранилища для сырых данных. *HBase* легко интегрируется с *Hadoop* для выполнения пакетной обработки больших данных. *HBase* обеспечивает высокопроизводительное хранение ключ–значение, что идеально подходит для хранения сырых данных о транспортных средствах. Эти данные включают статус транспортного средства, траектории движения, данные о действиях водителя и данные о заказах пользователей. В процессе принятия технических решений мы провели эксперименты по управлению сырыми данными о поездках с использованием СУБД, таких как *MySQL* и *PostgreSQL*.

Обзор существующих работ и технологий в области IoV. Работы в области *IoV* становятся всё более важными как в промышленности, так и в академических кругах. *IoV* является интегрированной технологией, включающей обработку больших данных, распределенные системы, управление большими данными, беспроводную связь, технологию конструирования транспортных средств, анализ человеческого поведения и другие аспекты. В облаке особенно критически важны обработка и управление большими данными.

Были предложены различные высокопроизводительные платформы для обработки больших данных, такие как *Storm*, *Spark*, *Samza*, *S4*, *Flume* и *Flink*. Однако по–прежнему сложно плавно интегрировать различные технологии для разработки систем для сложных сценариев в *IoV*.

Компании, предоставляющие услуги частного автотранспорта, такие как *Uber* и *Lyft*, разрабатывают сложные системы для обработки данных, интегрируя онлайн– и офлайн–обработку и управление большими данными для создания экосистемы для интеллектуальных транспортных средств.

Дизайн системы для потоковой обработки является сложной задачей исследования в *IoV*. На *Google* разработана модель потоковых данных для обработки неограниченных, неупорядоченных потоков данных. *Twitter* использует *Storm* для обработки потоковых данных. Компания *Amazon* предлагает *Kinesis* в качестве платформы для обработки больших данных. *Facebook* разрабатывает систему обработки данных в реальном времени. Предложены также различные системы для обработки и анализа больших пространственных данных, такие как *TrafficDB* и *Hadoop–GIS*.

Заключение. В данной статье описан опыт и решение проблем при создании *CarStream* – промышленной системы обработки больших данных для *IoV* и опыт построения нескольких приложений на основе этой системы для услуг автомобильного обслуживания с водителем. *CarStream* обеспечивает высокую надежность для сервисов, критичных для безопасности в *IoV*, включая трехуровневую систему мониторинга, охватывающую все слои от прикладного уровня до инфраструктурного. *CarStream* дополнительно использует кэширование в памяти и потоковую обработку для решения проблем обработки в реальном времени, обработки большого объема данных, низкого качества данных и низкой плотности информации. *CarStream* управляет большим объемом данных о поездках с помощью гетерогенной системы хранения данных. На

данный момент система может обрабатывать десятки тысяч автомобилей. В будущей работе планируется развивать *CarStream* в систему с архитектурой микросервисов для лучшего обслуживания системы и разработки новых приложений при увеличении числа автомобилей.

Список литературы

- [1] R. Ranjan. Streaming big data processing in datacenter clouds. *IEEE Cloud Computing*, 1(1):78–83, 2014.
- [2] M. Stonebraker, U. C. etintemel, and S. Zdonik. The 8 requirements of real-time stream processing. *ACM SIGMOD Record*, 34(4):42–47, 2005.
- [3] M. Tang, Y. Yu, Q. M. Malluhi, M. Ouzzani, and W. G. Aref. LocationSpark: A distributed in-memory data management system for big spatial data. In *Proc. VLDB Endow.*, 9(13):1565–1568, 2016.

Авторский вклад

Мигалевич Сергей Александрович – руководство исследованием потенциала современных информационных технологий в автомобильной промышленности.

Марков Алексей Николаевич – рассмотрение возможности совместного использования различных информационных технологий, технологий больших данных в автомобильной промышленности.

Ершов Денис Геннадьевич – исследование промышленной системы обработки больших данных для услуг аренды автомобилей с водителем.

INTEGRATION OF BIG DATA PROCESSING IN THE AUTOMOTIVE INDUSTRY

S.A. Migalevich
*Master of Technical Sciences,
Head of the Center for
Informatization and Innovative
Developments*

A.N. Markov
*Master of Technical Sciences,
Deputy Head of the Center for
Informatization and Innovative
Developments*

D.G. Ershov
*Student of the Belarusian State
University of Informatics and
Radioelectronics*

Abstract. As the Internet-of-Vehicles (IoV) technology becomes an increasingly important trend for future transportation, designing large-scale IoV systems has become a critical task that aims to process big data uploaded by fleet vehicles and to provide data-driven services. The IoV data, especially high-frequency vehicle statuses (e.g., location, engine parameters), are characterized as large volume with a low density of value and low data quality. Such characteristics pose challenges for developing real-time applications based on such data. In this paper, we address the challenges in designing a scalable IoV system by describing *CarStream*, an industrial system of big data processing for chauffeured car services. Connected with over 30,000 vehicles, *CarStream* collects and processes multiple types of driving data including vehicle status, driver activity, and passenger-trip information. Multiple services are provided based on the collected data. *CarStream* has been deployed and maintained for three years in industrial usage, collecting over 40 terabytes of driving data. This paper shares our experiences on designing *CarStream* based on large-scale driving-data streams, and the lessons learned from the process of addressing the challenges in designing and maintaining *CarStream*.

Keywords: Big Data, knowledge management, automotive industry, Generic awareness.

УДК 621.365.5

МОДЕЛИРОВАНИЕ ЛОКАЛЬНОГО ИНДУКЦИОННОГО НАГРЕВА ДЛЯ BGA И FLIP-CHIP МОНТАЖА



А.Д. Хацкевич

Магистр технических наук.

Инженер-электроник кафедры ЭТТ



В.Л. Ланин

Профессор кафедры электронной техники и технологии, доктор технических наук

А.Д. Хацкевич

Магистр технических наук. Окончил аспирантуру на кафедре ЭТТ. Инженер-исследователь. Имеет более 20 публикаций по данной тематике.

В.Л. Ланин

Профессор кафедры электронной техники и технологии. Имеет 30-летний опыт работы в области технологии сборки и монтажа электронной аппаратуры. Автор 10 монографий в данной области.

Анотация. Моделирование процесса индукционного нагрева шариков припоя при монтаже электронных модулей с *BGA* и *Flip-Chip* выполнено в программном пакете *COMSOL Multiphysics*. Выявлены закономерности воздействия концентратора и магнитопровода на обрабатываемый образец. Получены тепловые профили нагрева при различных частотах индукционного нагрева.

Ключевые слова: индукционный нагрев, моделирование, *BGA*, *Flip-Chip*.

Введение. Тенденция к уменьшению размеров и веса изделий в современной электронике, особенно в области высокочастотных устройств связанных с быстрым развитием телекоммуникационных технологий и аэрокосмической техники, приводит к разработке новых технологий, включая «многоуровневые» многокристальные модули. Процесс сборки электронных модулей с применением технологий *2,5D* и *3D* требует не только использования сложного технологического оборудования, но также соответствующих материалов, включая припойные материалы с разной температурой плавления для соединения трехмерной конструкции [1].

Один из методов монтажа кристаллов на платы – это метод *Flip-Chip*, при котором кристалл микросхемы размещается на выводах, выполненных непосредственно на его контактных площадках, расположенных по всей поверхности кристалла микросхемы [2].

Оплавление шариков припоя конвекцией горячего воздуха или в инфракрасной печи может привести к их перегреву в течении длительного времени, а также вызвать коробление печатной платы.

Индукционный нагрев является наиболее перспективным методом нагрева, представляющим собой бесконтактный способ передачи электрической энергии через токи высокой частоты. Этот метод включает в себя нагрев материалов, способных проводить электрическую энергию, при помощи токов высокой частоты, индуцированных переменным магнитным полем, то есть нагрев изделий магнитным полем индуктора [3].

Индукционный нагрев в микроэлектронике предоставляет несколько преимуществ перед другими методами пайки:

1 Индукционный нагрев осуществляется без прямого контакта с нагреваемым объектом что исключает возможность механического повреждения чувствительных компонентов, таких как микрочипы.

2 Локальный и точный нагрев в необходимой области позволяет избежать повреждения окружающих компонентов и минимизировать влияние на соседние элементы.

3 Энергия передается непосредственно в нагреваемый материал посредством индукции, что обеспечивает высокую эффективность процесса. Это может быть особенно важно при работе с элементами, где необходимо минимизировать потери энергии.

4 Индукционный нагрев позволяет легко управлять температурой в процессе оплавления шариков припоя, что важно для предотвращения перегрева и повреждения чувствительных компонентов.

5 Обеспечивает высокую скорость нагрева, что важно в микроэлектронике, где требуется быстрое и точное соединение компонентов.

Для моделирования процесса индукционной пайки был выбран программный пакет *COMSOL Multiphysics* как инструмент для анализа, решения и моделирования методом конечных элементов в различных физических и инженерных приложениях [4].

Создание и задание параметров модели. В качестве физической модели был выбран модуль *Induction Heating* в состав которого входят модули *Magnet Fields* и *Heat Transfer in Solids*. По умолчанию подключается модуль *Electromagnetic Heating*. Он необходим при моделировании задач, связанных с электротермическим нагревом, таких как индукционный, лазерный или микроволновый нагрев. В нашем случае свойства материалов зависят от температуры, поэтому моделирование параметров электромагнитного поля будет меняться по мере нагрева системы. Геометрическая модель локального индукционного нагрева представлена на рисунке 1.

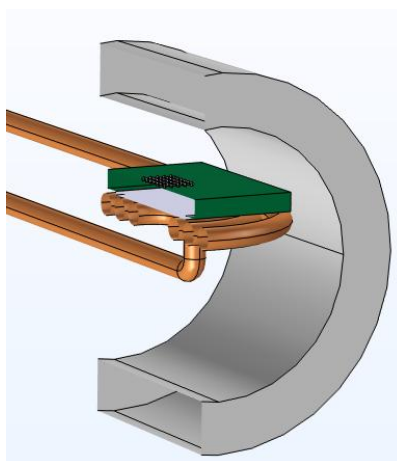


Рисунок 1. Модель локального индукционного нагрева

Параметры компонентов геометрической модели индукционного нагрева представлены в таблице 1.

Таблица 1. Геометрические параметры компонентов

Геометрические размеры платы	25x25x2.54 мм
Геометрические размеры концентратора	20x20x3 мм
Диаметр трубки индуктора	3 мм
Размер шариков припоя	0.760 мм

Для повышения эффективности индукционного нагрева в качестве концентратора выбран медная пластина с симметричным паттерном прямоугольных окон, которые необходимы для предотвращения перегрева печатной платы, как происходит в случае использования концентратора без отверстий. Внешний вид концентратора представлен на рисунке 2.

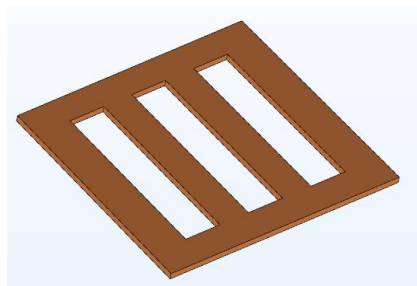


Рисунок 2. Внешний вид концентратора

Материалы и их свойства, использованные в модели, взяты из стандартной библиотеки материалов *COMSOL* и представлены в таблице 2. В качестве шариков припоя выбран бессвинцовый припой, что соответствует современным экологическим стандартам в электронике.

Таблица 2. Материалы модели

Шарики припоя	$Sn-3.5 Ag-0.5 Cu$
Материал индуктора	медь
Материал концентратора	медь
Материал печатной платы	FR-4
Магнитопровод	Феррит с магнитной проницаемостью 2500

Построение сетки представляет собой важный этап в моделировании, так как он напрямую определяет затраты вычислительных ресурсов при расчете. Более того, создание сетки – один из наиболее ресурсоёмких этапов при настройке и расчёте модели методом конечных элементов. *COMSOL Multiphysics* предоставляет возможность настройки сетки как в автоматическом, так и в ручном режимах. Однако, в контексте данной задачи, автоматический режим не является подходящим, поскольку из-за маленького размера шариков припоя, сетка становится неоптимальной. Вместо этого, мы используем размер сетки *Extra Fine* для шариков припоя, а для остальных элементов применяем сетку большего размера. Это позволяет оптимизировать модель и значительно сократить время её расчёта. Внешний вид настроенной сетки в модели локального индукционного нагрева представлен на рисунке 3.

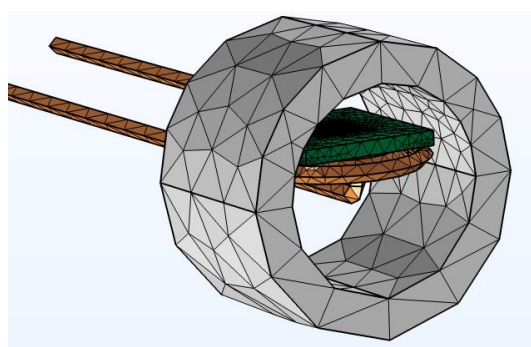


Рисунок 3. Настроенная сетка модели локального индукционного нагрева

Выбор решателя. *Coil Geometry Analysis* в *COMSOL* используется для расчета тока в функции *Coil* в 3D-моделях.

Этот анализ используется для определения локального направления тока в области, учитывая входные и выходные границы катушки (или внутреннюю границу в случае катушек замкнутого типа). Это особенно полезно, когда модель катушки зависит от выбранной модели проводника. Условия границ для анализа геометрии указываются с помощью подузлов ввода и вывода, доступных с узлом. Это обеспечивает более точное и эффективное моделирование электромагнитных систем.

Для расчета в частотной и временной области используется *Frequency Domain* и *Time Dependent* соответственно. Решатели в интерфейсе программы представлены на рисунке 4.

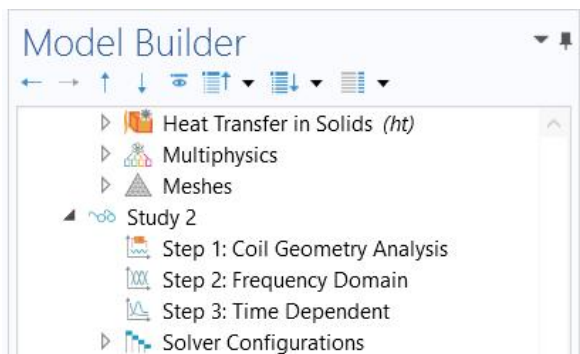


Рисунок 4. Последовательность решателей в узле *Study*

Температурные поля платы на выводах в виде шариков припоя, выполненных на контактных площадках платы представлены на рисунке 5.

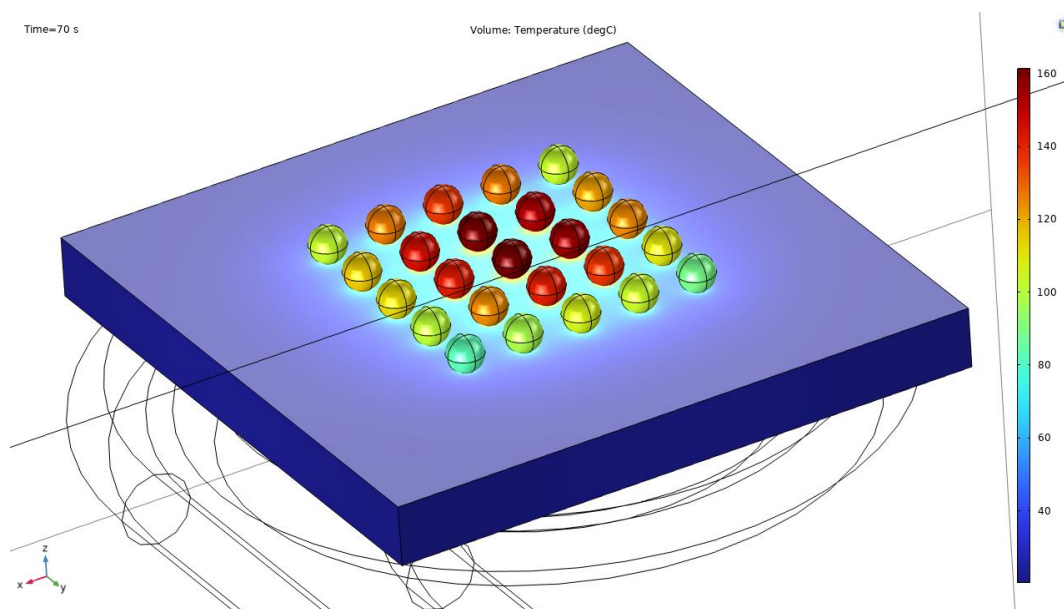


Рисунок 5. Температурные поля в шариках припоя

По равномерности поля можно сделать вывод, что на равномерность нагрева шариков припоя влияет геометрическая форма индуктора. Краевые шарики нагреваются слабее на 10 - 15 °C в следствии краевого эффекта.

Термопрофили локального индукционного нагрева на различных частотах без концентратора и магнитопровода представлен на рисунке 6.

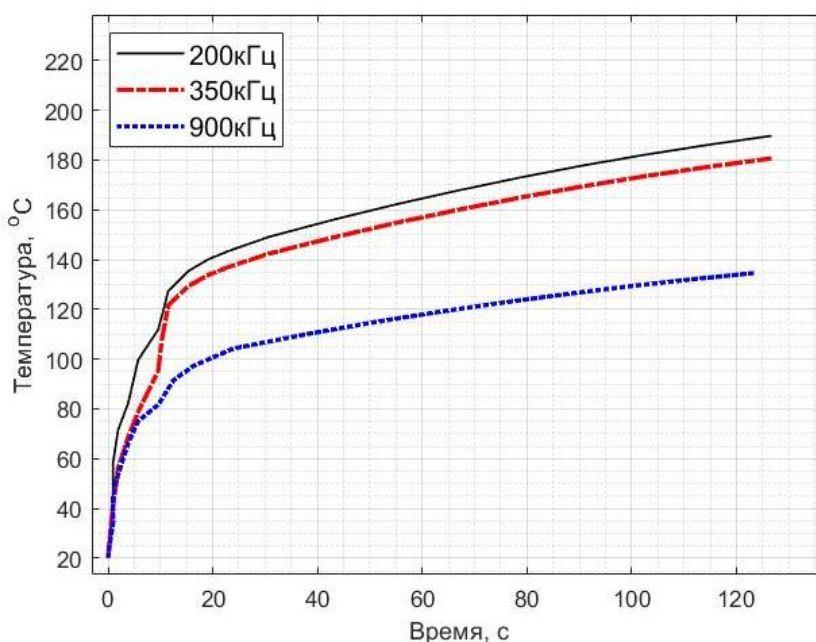


Рисунок 6. Термопрофили локального индукционного нагрева на различных частотах без концентратора и магнитопровода

Температура нагрева шариков припоя без концентратора и магнитопровода не достигает необходимых значений для пайки бессвинцовых шариков припоя на всех частотах индукционного нагрева.

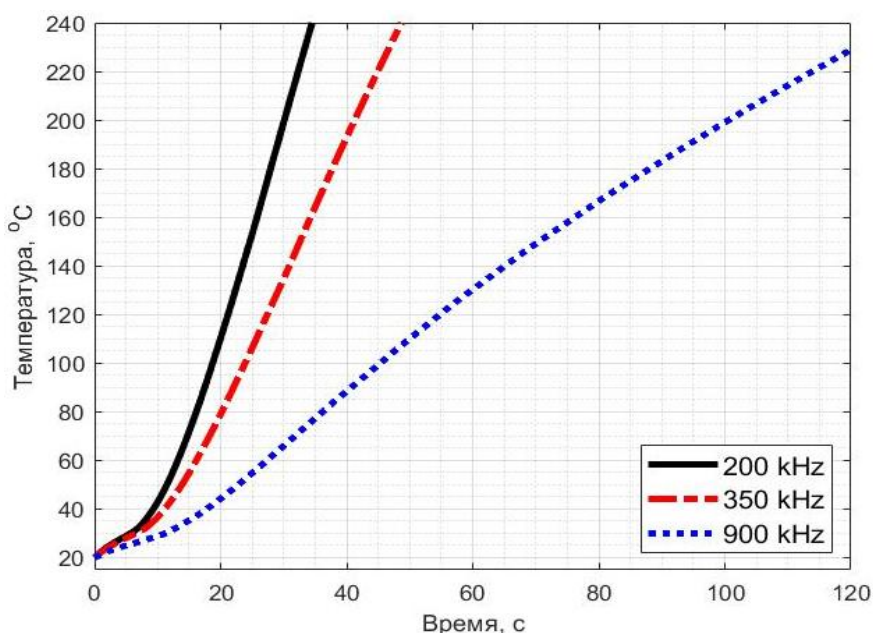


Рисунок 7. Термопрофили локального индукционного нагрева на различных частотах с концентратором и магнитопроводом

Анализируя графики на рисунке 7, можно заметить, что применение концентратора и магнитопровода увеличило скорость нагрева: для частоты 200 кГц она составила 7–8°C/с, для 350 кГц – 4–5°C/с, для 900 кГц – 2–3°C/с.

Использование частот меньше 200 кГц не рекомендуется, поскольку это может привести к перегреву медного концентратора, что, в свою очередь, может вызвать перегрев и последующее разрушение печатной платы. В связи с этим, для данной конфигурации паяемого образца и мощности нагрева (30 Вт), рабочий диапазон частот следует выбирать в пределах 400 кГц – 1 МГц.

Заключение. Процесс индукционного нагрева оптимизирован с помощью программы *COMSOL Multiphysics*. Показано, что применение медного концентратора и ферритового магнитопровода повышает эффективность нагрева до 5 раз. Однако частота индуктора существенно влияет на температуру паяемого образца. В данном случае частота 200 кГц приводит к чрезмерно быстрому нагреву шариковых выводов, что негативно сказывается на качестве такого соединения. В данной модели при мощности индукционной системы в 30 Вт рабочий диапазон частот следует выбирать в пределах 400 кГц – 1 МГц.

Список литературы

- [1] Технологии и оборудование субмикронной электроники / А.П. Достанко [и др.]; под общ. ред. акад. А.П. Достанко. – Минск: Беларуская навука, 2020.–260с.
- [2] Flip chip – [Электронный ресурс]. – Режим доступа: https://semiengineering.com/knowledge_centers/packaging/flip-chip/. – Дата доступа: 23.01.2024.
- [3] Индукционный нагрев – [Электронный ресурс]. – Режим доступа: <https://www.элсит.рф/вики/термообработка>. – Дата доступа: 10.04.2023.
- [4] COMSOL Multiphysics – [Электронный ресурс]. – Режим доступа: <https://www.comsol.ru/comsol-multiphysics>. – Дата доступа: 10.04.2023.

Авторский вклад

Авторы внесли равноценный вклад в написании статьи

SIMULATION OF LOCAL INDUCTION HEATING FOR BGA AND FLIP-CHIP ASSEMBLY

A.D. Khatskevich

Master of Engineering.

PhD applicant. *Electronic Engineer of the Department of ETT*

V.L. Lanin

Professor, Department of Electronic System and Technology, Doctor of Sciences

Abstract. The process of induction heating of solder balls using simulation in the COMSOL Multiphysics software package has been investigated. The patterns of the concentrator and magnetic circuit's impact on the processed sample have been identified. Heat profiles of heating at various inductor frequencies have been obtained.

Keywords: induction heating, simulation, BGA, Flip-Chip.

УДК 004.522

ОБРАБОТКА БОЛЬШИХ ДАННЫХ



Аннаныязова Г.

Преподаватель кафедры «Программное обеспечение информационных технологий»
Институт Телекоммуникаций
и информатики Туркменистана,
g.annanyazova@gmail.com

Аннаныязова Г.

Преподаватель кафедры «Программное обеспечение информационных технологий» Институт Телекоммуникаций и информатики Туркменистана.

Аннотация. Технология обработки *Big Data* сводится к трем основным направлениям, которые решают три типа задач, а именно, перевод и хранение поступающей информации в гигабайтах, терабайтах, петабайтах и т.д. данных для их обработки, хранения и применения на практике; структурирование разрозненного контента: фотографий, текстов, аудио, видео и всех других видов данных; анализ Больших данных и внедрение разных методов обработки неструктурированных данных, создание разных аналитических отчетов. Чтобы упростить и ускорить работу с большими данными есть много функций языков программирования. На данный момент библиотека *Pandas* является ключевой в анализе данных (*Data Mining*). *Pandas* – это быстрый, мощный, гибкий и простой в использовании инструмент для анализа и обработки данных с открытым исходным кодом, созданный на языке программирования *Python*.

Ключевые слова: Большие данные, обработка, прием данных, сбор данных, анализ данных, представление результатов, *Python*, *Pandas*.

Введение. *Big Data* или большие данные – это структурированные или неструктурированные массивы данных большого объема. Их обрабатывают при помощи специальных автоматизированных инструментов, чтобы использовать для статистики, анализа, прогнозов и принятия решений.

Сам термин «большие данные» предложил редактор журнала *Nature* Клиффорд Линч в спецвыпуске 2008 года. Он говорил о взрывном росте объемов информации в мире. К большим данным Линч отнес любые массивы неоднородных данных более 150 Гб в сутки, однако единого критерия до сих пор не существует. До 2011 года анализом больших данных занимались только в рамках научных и статистических исследований. Но к началу 2012-го объемы данных выросли до огромных масштабов, и возникла потребность в их систематизации и практическом применении.

Большие данные, в свою очередь, становятся источником новых данных, которые стимулируют аналитические инновации в бизнесе, правительстве и академических кругах. Эти нововведения в состоянии радикально изменить взгляд организаций на свой бизнес. Большие данные обеспечат информацию, которая поможет принимать более взвешенные решения, и в некоторых случаях они будут разительно отличаться от тех, что принимаются сегодня. Анализ больших данных даст такое понимание, о котором сегодня

можно только мечтать. Вы увидите, что укрощение волны больших данных и укрощение новых источников данных осуществляется аналогичными способами. Тем не менее дополнительные возможности, которые предоставляют большие данные, требуют использования новейших инструментов, технологий, методов и процессов. [1]

Обработка больших данных. Обработка больших данных в настоящее время с помощью обычных программных методов и аппаратных средств зачастую невозможна, так как этого не позволяет огромный объем имеющейся информации. Несмотря на это, с учетом своих сложностей, несколько эффективных методов для проведения подобных операций все же существует. Для каждого конкретного случая необходимо выбирать наиболее подходящий способ обработки данных, только тогда результаты окажутся удовлетворительными: и с технологической, и с экономической точки зрения.

Если же говорить о терминологии, то «*Big Data*» подразумевает не только данные как таковые, но и принципы обработки больших данных, возможность дальнейшего их использования, порядок обнаружения конкретного информационного блока в больших массивах. Вопросы, связанные с такими процессами, не теряют своей актуальности. Их решение носит важный характер для тех систем, которые многие годы генерировали и копили различную информацию.

Существуют критерии информации, определенные в 2001 году *Meta Group*, которые позволяют оценить, соответствуют ли данные понятию *Big Data* или нет:

- **Volume (объем)** – примерно 1 Петабайт и выше.
- **Velocity (скорость)** – генерация, поступление и обработка данных с высокой скоростью.
- **Variety (разнообразие)** – разнородность данных, различные форматы и возможное отсутствие структурированности.

Архитектура системы обработки. В этом разделе мы будем рассматривать архитектуру системы обработки Больших данных. Для работы с Большими данными используются сложные системы, в которых можно выделить несколько компонентов или слоёв (Layers). Обычно выделяют четыре уровня компонентов таких систем: прием, сбор, анализ данных и представление результатов (рисунок 1).



Рисунок 1. Стек работы с Большими данными

Это деление является в значительной мере условным так как, с одной стороны, каждый компонент в свою очередь может быть разделен на подкомпоненты, а с другой некоторые функции компонентов могут перераспределяться в зависимости от решаемой

задачи и используемого программного обеспечения, например, выделяют хранение данных в отдельный слой.

Для работы с Большими данными разработчиками систем создаются модели данных, содержательно связанные с реальным миром. Разработка адекватных моделей данных представляет собой сложную аналитическую задачу, выполняемую системными архитекторами и аналитиками. Модель данных позволяет создать математическую модель взаимодействий объектов реального мира и включает в себя описание структуры данных, методы манипуляции данными и аспекты сохранения целостности данных. Описание разработки моделей данных не является задачей настоящего руководства.

Для хранения данных используются распределенные системы различных типов. Это могут быть файловые системы, базы данных, журналы, механизмы доступа к общей виртуальной памяти. Большинство систем хранения ориентированы исключительно на работу с Большими данными, они имеют крайне ограниченное число функций (например, может отсутствовать возможность не только модификации, но и удаления поступивших данных) что объясняется внутренней сложностью создания высокоэффективных распределенных систем.

Для того, чтобы работа с данными происходила быстрее системы хранения и обработки данных распараллеливаются в кластере (*cluster*, группа компьютеров, объединенных сетью для выполнения единой задачи). Однако, согласно гипотезе Брюера невозможно обеспечить одновременную согласованность (непротиворечивость) данных, доступность данных и устойчивость системы к отделению отдельных узлов. Гипотеза доказана для транзакций типа *ACID* (*Atomic, Consistent, Isolated, Durable*) и известна под названием *CAP* теоремы (*Consistency, Availability, Partition tolerance*).

Прием данных (*Data Ingestion*). Источники данных имеют различные параметры, такие как частоту поступления данных из источника, объем порции данных, скорость передачи данных, тип поступающих данных и их достоверность.

Для эффективного сбора данных необходимо установить источники данных. Это могут быть хранилища данных, поставщики агрегированных данных, *API* каких-либо датчиков, системные журналы, сгенерированный человеком контент в социальных сетях, в корпоративных информационных системах, геофизическая информация, научная информация, унаследованные данные из других систем.

Источники данных определяют исходный формат данных. Например, мы можем самостоятельно проводить погодные на территории аэропорта, использовать данные, поступающие с взлетающих и садящихся самолетов, закупить данные со спутников, пролетающих над аэропортом и у местной метеослужбы, а также найти их где-то в сети в другом месте. В общем случае для каждого источника необходимо создавать собственный сборщик (*Data Crawler* для сбора информации в сети и *Data Acquisition* для проведения измерений).

Прием данных заключается в начальной подготовке данных от источников с целью приведения данных к общему формату представления данных. Этот единый формат выбирается в соответствии с принятой моделью данных. Выполняются преобразования систем измерения, типов (типизация), верификация. Обработка данных содержательно не затрагивает имеющуюся в данных информацию, но может изменять ее представление (например, приводить координаты к единой системе координат, а значения к единой размерности).

Сбор данных (*Data Staging*). Этап сбора данных характеризуется непосредственным взаимодействием с системами хранения данных. Устанавливается точка сбора, в которой собранные данные снабжаются локальными метаданными и помещаются в хранилище либо передаются для последующей обработки. Данные, по каким-либо причинам не прошедшие точку сбора, игнорируются.

Для структурированных данных проводится преобразование из исходного формата по заранее заданным алгоритмам. Это наиболее эффективная процедура в случае, если структура данных известна. Однако если данные представлены в двоичном виде, структура и связи между данными утеряны, то разработка алгоритмов и основанного на них программного обеспечения для обработки данных может оказаться крайне затруднительной.

Для полуструктурированных данных требуется интерпретация поступающих данных и использование программного обеспечения, умеющего работать с используемым языком описания данных. Существенным плюсом полуструктурированных данных является то, что в них зачастую содержатся не только сами данные, но метаданные в виде информации о связях между данными и способах их получения.

Разработка программного обеспечения для обработки полуструктурированных данных представляет собой достаточно сложную задачу. Однако имеется значительное количество готовых конвертеров, которые могут, например, извлечь данные из формата *XML* в сформированное табличное представление.

Наибольшего объема работ требует обработка неструктурированных данных. Для их перевода к заданному формату может потребоваться создание специального ПО, сложная ручная обработка, распознавание и выборочный ручной контроль.

На этапе сбора проводится контроль типов данных и может выполняться базовый контроль достоверности данных. Например, координаты молекул газа, содержащихся в какой-либо области, не могут лежать за пределами этой области, а скорости – существенно превышать скорость звука. Для того, чтобы избежать ошибок типизации, необходимо проверять правильно ли заданы единицы измерения. Например, в одном наборе данных высота может измеряться в километрах, а в другом – в футах. В этом случае необходимо произвести преобразование высоты в те единицы измерения, которые приняты в используемой модели.

При сборе данные систематизируются и снабжаются метаданными, хранимыми в связанных метаданных. При наличии большого количества источников данных может потребоваться управление сбором данных для того, чтобы сбалансировать объемы информации, поступающие из различных источников.

Собранные данные либо сохраняются в системах хранения, либо (в особенности, для потоковых данных) передаются для анализа в реальном времени.

Анализ данных (*Analysis Layer*). Анализ данных, в отличие от сбора данных, использует информацию, содержащуюся в самих данных. Анализ может проводиться как в реальном времени, так и в пакетном режиме. Анализ данных составляет основную по трудоемкости задачу при работе с Большими данными.

Существует множество методик обработки данных: предиктивный анализ, запросы и отчетность, реконструкция по математической модели, трансляция, аналитическая обработка и другие. Методики используют специфические алгоритмы в зависимости от поставленных целей. Например, аналитическая обработка может являться анализом изображений, социальных сетей, географического местоположения, распознавания по признакам, текстовым анализом, статистической обработкой, анализом голоса, транскрибированием.

Алгоритмы анализа данных также, как и алгоритмы обработки данных, опираются на модель данных. При этом при анализе может быть использовано несколько моделей, задающих общий формат данных, но по-разному моделирующие содержательные процессы, данные о которых мы обрабатываем.

При использовании при анализе методов искусственного интеллекта, в частности нейронных сетей, производится динамическое обучение моделей на различных наборах данных.

При анализе данных производится идентификация сущностей, описываемых данными на основании имеющейся в данных информации и используемых моделей.

Сущностью анализа является аналитических механизм, использующий аналитические алгоритмы, управление моделями и идентификацию сущностей для получения новой содержательной информации, являющейся результатом анализа.

Представление результатов (*Consumption Layer*). Результаты анализа данных предоставляются на уровне потребления. Имеется несколько механизмов, позволяющих использовать результаты анализа больших данных.

1 Мониторинг метаинформации.

Подсистема отображения в реальном времени существенных параметров работы системы, загруженности вычислителей, распределение задач в кластере, распределение информации в хранилищах, наличие свободного места в хранилищах, поступление данных от источников, активности пользователей, отказов оборудования и тд.

2 Мониторинг данных.

Подсистема отображения в реальном времени процессов приема, сбора и анализа данных, навигация по данным.

3 Генерация отчетов, запросы к данным, представление данных в виде визуализации на дэшбордах (*Dashboard*), в формате *PDF*, инфографике, сводных таблицах и кратких справках.

Обработка данных. Чтобы упростить и ускорить работу с большими данными есть много функций языков программирование. Один из первых инструментов, с которым сталкивается аналитик либо *Data Scientist* – это *Pandas*, библиотека *Python* для обработки и анализа данных.

Язык *Python* помогает упростить анализ данных. Если вы научились пользоваться электронными таблицами, то сможете освоить и *pandas*. Несмотря на сходство с табличной компоновкой *Excel*, *pandas* обладает большей гибкостью и более широкими возможностями. Эта библиотека для *Python* быстро выполняет операции с миллионами строк и способна взаимодействовать с другими инструментами. Она дает идеальную возможность выйти на новый уровень анализа данных. С её помощью мы импортируем и сортируем данные, делаем выборки и находим зависимости. Например, чтобы прочитать файл средствами *Pandas* в *Python* мы пишем:

```
import pandas as pd
data = pd.read_csv('data_file.csv')
```

Такой подход простой и понятный. Каждый *Data Scientist* или аналитик знает это. Но если данных много? Скажем 100 000 000 строк, они постоянно меняются, сроки горят, и до обеда надо проверить еще 100 гипотез?

Возьмем исследовательский набор данных о диабете с сайта *Kaggle* и продублируем каждую строку 100 000 для создания нашего тестового набора данных. В результате получится 76 800 000 строк.

Изначальный датасет выглядит так:

Таблица 1. Набор данных о диабете.

	<i>Pregnancies</i>	<i>Glucose</i>	<i>Blood Pressure</i>	<i>Skin Thickness</i>	<i>Insulin</i>	<i>BMI</i>	<i>Diabetes Pedigree Function</i>	<i>Age</i>	<i>Outcome</i>
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1

Окончание таблицы 1

	<i>Pregnancies</i>	<i>Glucose</i>	<i>Blood Pressure</i>	<i>Skin Thickness</i>	<i>Insulin</i>	<i>BMI</i>	<i>Diabetes Pedigree Function</i>	<i>Age</i>	<i>Outcome</i>
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
...
763	10	101	76	48	180	32.9	0.171	63	0
764	2	122	70	27	0	36.8	0.340	27	0
765	5	121	72	23	112	26.2	0.245	30	0
766	1	126	60	0	0	30.1	0.349	47	1
767	1	93	70	31	0	30.4	0.315	23	0

Преобразовываем его для наших экспериментальных задач:

```
# Чтение датасета из файла
df = pd.read_csv('diabetes.csv')

# Создание тестового файла
df = df.loc[df.index.repeat(100000)]

# Сохранение в файл для экспериментов
df.to_csv('benchmark.csv')

print(len(df), 'строков')
76 800 000 строк
```

Преобразование было долгим, но будем надеется, что время потрачено не зря. Сколько же займет чтение такого файла?

```
df = pd.read_csv('benchmark.csv')
```

Если каждый раз загружать изменившийся датасет заново, на наши 100 гипотез уйдет 13 часов. Сделать до обеда не получится.

Мы хотим загружать данные быстрее, но при этом не терять всех преимуществ, которые даёт *Pandas*. Простая функция использующая *datatable* поможет нам сделать это:

```
import datatable as dt
import pandas as pd
def read_fast_csv(f):
    frame = dt.fread(f)
    ds = frame.to_pandas()
    return ds
```

Попробуем теперь прочитать наш большой датасет:

```
ds = read_fast_csv('benchmark.csv')
```

Это в 6 раз быстрее!

Столбцы с датами в разных форматах. Для анализа и визуализации данных часто требуется работать с датами в разных форматах, когда дата дана только одна. А еще хуже, если все даты прописаны как попало. Приходится приводить все даты к стандартному виду, а потом программным путем добывать из каждой даты разную информацию. Упрощает этот процесс сначала смоделируем ситуацию и создадим функцию, и к нашему

исходному датасету добавим колонку со случайными датами в диапазоне двух последних лет:

```
import numpy as np

# Чтение датасета из файла
df = pd.read_csv('diabetes.csv')
df['date'] = np.random.choice(pd.date_range('2020-01-01', '2022-12-31'), 768)
df.to_csv('diabetes_dates.csv')
```

Ниже код функции, которая к существующему датасету добавляет колонки с наиболее частыми требуемыми форматами, чтобы сделать выборки по годам, по кварталам, месяцам или неделям, либо взять полную дату для целей визуализации.

```
import datetime
from datetime import timedelta

def granular_dates(df, col):
    df['ts_date'] = pd.to_datetime(df[col]).dt.normalize()
    # Полная дата с названием месяца
    df['ts_date_str'] = df["ts_date"].dt.strftime("%d %B %Y")
    # Краткая дата с сокращением месяца
    df['ts_date_str_short'] = df["ts_date"].dt.strftime("%d %b %Y")
    # Только год
    df['ts_year'] = df.ts_date.dt.year
    # Только номер месяца
    df['ts_month'] = df.ts_date.dt.month
    # Только число
    df['ts_day'] = df.ts_date.dt.day
    # Год и квартал
    df['ts_quarter'] = pd.PeriodIndex(df.ts_date, freq="Q").astype(str)
    # Номер недели
    df['ts_dayweek'] = df.ts_date.dt.dayofweek
    # День недели
    df['ts_dayweektext'] = df.ts_date.dt.strftime("%a")
    # Дата конца недели (воскресенья)
    df['ts_week_start'] = df.ts_date.apply(lambda x: x -
timedelta(days=x.weekday())).dt.strftime("%b-%d")
    # Дата конца недели (воскресенья)
    df['ts_week_end'] = df.ts_date.apply(lambda x: x -
timedelta(days=x.weekday()) + timedelta(days=6)).dt.strftime("%b-%d")
```

Теперь всего одна строчка кода (не считая чтение файла):

```
# Чтение датасета
df = pd.read_csv('diabetes_dates.csv')
# Добавление колонок с датами разного формата
granular_dates(df, 'date')
```

Кроме обработке больших данных с помощью *Python* пакета *Pandas*, но также можно достичь этой функциональности при помощи нескольких генераторов. Для обработки больших наборов данных или потоков данных без максимального использования памяти компьютера **создаются конвейеры данных, построенного на генераторах** позволяющие скомпоновать код. Представим, что есть большой *CSV*-файл *big-data.csv* в несколько тысяч строк с данными посещения сайта, которые нужно обработать.

Для обработки такого файла необходимо проделать следующие действия:

- 1 Прочитать каждую строчку файла.

- 2 Разбить каждую строку на список значений.
- 3 Извлечь имена столбцов.
- 4 Использовать имена столбцов и список значений из строк для создания словарей.
- 5 Отфильтровать значения, которые не интересуют.
- 6 Обработать интересные значения.

Начнем обработку больших данных с чтения каждой строки из *CSV*-файла с помощью выражения генератора. Генераторы ленивы, следовательно, не загружают файл целиком в оперативную память, а читают его построчно.

```
file_name = "big-data.csv"
lines = (line for line in open(file_name))
```

Здесь чтение *CSV*-файла при помощи выражения генератора показано в качестве примера и понимания конвейеров. Для работы с *CSV*-файлами существует специальный оптимизированный модуль *csv*, который включен в стандартную библиотеку *Python*.

Затем будем использовать другое выражение-генератор совместно с предыдущим, чтобы разбить каждую строку на список. Создадим генератор *list_line*, который выполняет итерацию по строкам генератора *lines*, который в свою очередь будет построчно читать большой файл.

```
list_line = (s.rstrip().split(",") for s in lines)
```

По сути - один генератор вкладывается в другой. Это общий шаблон, который используют при проектировании конвейеров генераторов.

Примечание: Код использует метод строки *str.rstrip()* в выражении генератора *list_line*, чтобы убедиться в отсутствии завершающих символов новой строки, которые могут присутствовать в *CSV*-файлах.

Затем извлечем имена столбцов из *big-data.csv*. Так как имена столбцов обычно являются первой строкой *CSV*-файла, то их можно получить при помощи вызова встроенной функции *next()*:

```
cols = next(list_line)
```

Вызов функции *next()* сохранит список названий колонок в переменную *cols* и переместит итератор на следующую строку генератора *list_line*, с которой уже начинаются данные.

Чтобы иметь возможность фильтровать и выполнять операции с полученными данными, создадим словари, где ключами будут имена столбцов из *CSV*-файла:

```
log_dicts = (dict(zip(cols, data)) for data in list_line)
```

Выражение-генератора *log_dicts* перебирает списки, созданные генератором *list_line* и использует функции *zip()* и *dict()* для создания словаря, где ключами будут имена столбцов *cols*, а значения – соответствующие данные.

Для обработки данных нужно использовать четвертый генератор, который, например, будет фильтровать данные по столбцу *user_agent*, в котором есть вхождение строки «*YandexBot*» и вытаскивать соответствующие значения столбца *ip_address*:

```
finding = (
    log_dict["ip_address"]
    for log_dict in log_dicts
    if "YandexBot" in log_dict["user_agent"]
)
list_ip = list(
    finding
)
```

В этом фрагменте кода выражение-генератор *finding* перебирает результаты генератора *log_dicts* и возвращает значение ключа *ip_address* для любого словаря *log_dict*, в котором значение «*YandexBot*» встречается в значении словаря с ключом *user_agent*.

Код выше не перебирает все сразу в выражении генератора *finding*. На самом деле ничего не будет исполняться, пока не будет задействован цикл *for* или функция, которая работает с итерациями, например *list()*.

Проще говоря, вызов *list(finding)* заставляет работать все генераторы в коде.

И так, собираем код вместе:

```
file_name = "big-data.csv"
# выражение-генератор, который получает строки из файла
lines = (line for line in open(file_name))
# выражение-генератор, который получает список полей из каждой строки
list_line = (s.rstrip().split(",") for s in lines)
# получение списка имен колонок - это первая строка
cols = next(list_line)
# выражение-генератор, создающий словари
log_dicts = (dict(zip(cols, data)) for data in list_line)
# выражение-генератор, фильтрующий нужные значения
finding = (
    log_dict["ip_address"]
    for log_dict in log_dicts
    if "YandexBot" in log_dict["user_agent"]
)
# преобразование генератора в итоговый список
list_ip = list(finding)
```

Этот сценарий объединяет все созданные генераторы, и все они функционируют как один конвейер больших данных.

Заключение. В этой статье мы представили архитектуры обработки больших данных, в которых можно выделить несколько компонентов или слоёв (*Layers*). Детально описано четыре уровня компонентов систем: прием, сбор, анализ данных и представление результатов. А также, для упрощения и ускорение работы с большими данными и их обработки, привели несколько примеров используя библиотеки Pandas языка программирования Python и создание конвейера данных построенных на генераторах.

Список литературы

[1] Билл Фрэнкс. пер. с англ. Андрея Баранова. Угрожение больших данных: как извлекать смысл из гигантских потоков данных с помощью продвинутой аналитики. — М. : Манн, Иванов и Фербер, 2014.

[2] Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018.

PROCESSING BIG DATA

Annanyuzova G.

*Lecturer of the Department of Software of information technologies,
The Institute of Telecommunications and Informatics of Turkmenistan*

Annotation. Big Data processing technology comes down to three main areas that solve three types of problems, namely, translation and storage of incoming information in gigabytes, terabytes, petabytes, etc. data for their processing, storage and application in practice; structuring disparate content: photographs, texts, audio, video and all other types of data; analysis of Big Data and implementation of various methods for processing unstructured data, creation of various analytical reports. To simplify and speed up working with big data, there are many functions of programming languages. At the moment, the Pandas library is key in data analysis (Data Mining). Pandas is a fast, powerful, flexible, and easy-to-use open-source data science tool built in the Python programming language.

Keywords: Big data, processing, Data Ingestion, Data Staging, Analysis Layer Consumption Layer, Python, Pandas.

УДК 004.9:37.014.3

BIG DATA В ОБРАЗОВАНИИ: КАК ИСПОЛЬЗОВАНИЕ ДАННЫХ ПОМОГАЕТ ПЕРСОНАЛИЗИРОВАТЬ УЧЕБНЫЙ ПРОЦЕСС И ПОВЫШАТЬ УСПЕВАЕМОСТЬ В ШКОЛЕ



Д. В. Дудко
учащаяся ГУО «Лидский
РЦТТ»
diana12200709@gmail.com



М. В. Дашкевич
Педагог дополнительного
образования ГУО «Лидский РЦТТ»,
магистр экономических наук
martadashkevich22@gmail.com

Д. В. Дудко

Учащаяся ГУО «Лидский РЦТТ» объединения по интересам «Технологии в экономике» и «Программирование на Python», учащаяся 10 класса «Средней школы номер 16 имени П. М. Машерова города Лиды»

М. В. Дашкевич

Окончила УО «Гродненский государственный университет имени Янки Купалы». Область научных интересов - проект как одна из форм научно-исследовательской деятельности учащихся на занятиях объединений по интересам.

Аннотация. Данная статья исследует роль *Big Data* в современном образовании с фокусом на школьное обучение. Проанализированы способы использования данных для персонализации учебного процесса и улучшения успеваемости учащихся. Обсуждаются методы сбора и анализа информации о предпочтениях, потребностях и уровне знаний учащихся для создания индивидуализированных учебных планов. Представлены примеры использования *Big Data* для оптимизации расписания занятий, идентификации проблемных областей, прогнозирования успеха учащихся и повышения их мотивации. Рассмотрены практические примеры применения аналитики данных в школьном образовании с целью улучшения качества обучения и индивидуализации подхода к каждому учащемуся. Главный вывод состоит в том, что использование *Big Data* позволяет эффективно адаптировать учебный процесс к потребностям каждого ученика, что способствует повышению успеваемости и качества образования в школе.

Ключевые слова: *Big Data*, образование, учебный процесс, персонализация, успеваемость, школа, конфиденциальность, анализ данных, индивидуализация обучения, оптимизация расписания, мотивация учащихся.

Введение. Современное образование сталкивается с вызовами, требующими инновационных подходов для повышения эффективности обучения и улучшения результата каждого учащегося. В этом контексте *Big Data*, или большие массивы данных, представляют собой мощный инструмент, способный преобразить традиционный учебный процесс и сделать его более индивидуализированным и адаптированным к потребностям каждого учащегося. Анализ данных в образовании открывает новые возможности для

персонализации обучения через понимание уникальных потребностей каждого учащегося. Путем сбора, обработки и интерпретации информации о предпочтениях, успеваемости, стиле обучения, и других параметрах учащихся, школы могут создавать индивидуализированные учебные планы, оптимизировать расписание занятий, и повышать мотивацию учащихся к обучению. В данной статье рассматривается использование *Big Data* в образовании, способствуя персонализации учебного процесса и повышению успеваемости в школе.

Методы работы с BIG DATA. При сборе и анализе данных применяют различные методы:

1 **Машинное обучение и нейронные сети.** Использование в анализе нейронных сетей, которые обучаются вычислениям на массивах подготовленных данных. Обученная нейросеть может обрабатывать большие данные с большой точностью. Чтобы нейросеть работала, ее нужно сначала обучить – этот процесс называется машинным обучением. После обучения можно использовать нейросеть для обработки *Big Data*.

2 **Смешение и интеграция данных.** Приведение неструктурированных разнородных данных из разных источников к единому виду, например, текстовому.

3 **Data Mining.** Анализ данных для выявления закономерностей в больших неструктурированных массивах данных.

4 **Краудсорсинг.** Привлечение людей к анализу данных. В небольших разовых проектах *Big Data* найм людей или поиск волонтеров для проверки таблиц и баз данных может оказаться дешевле, чем компьютерные методы.

5 **Предиктивная аналитика.** Анализ больших данных за прошлые периоды для прогнозирования их поведения в будущем. Например, анализ успеваемости учащихся за прошлые годы позволит выявить слабые стороны в обучении и укрепить их.

6 **Имитационное моделирование.** Анализ данных с возможностью изменять параметры для гипотетических ситуаций, например, для прогнозирования успеваемости учащихся.

Использование BIG DATA в школе. Использование *Big Data* в образовании, включая школьное обучение, представляет значительный потенциал для улучшения образовательной системы, персонализации учебного процесса и повышения успеваемости учащихся.

Способы использования *Big Data* в школах:

1 **Персонализированное обучение:** анализ данных может помочь педагогам понять потребности каждого учащегося, их уровень знаний, стиль обучения и индивидуальные требования. На основе этих данных учителя могут создавать персонализированные учебные планы и материалы, адаптированные к потребностям каждого учащегося.

2 **Предсказание успеха учащихся:** анализ *Big Data* может помочь школам прогнозировать успеваемость учащихся, выявлять риски дропаутов и предпринимать меры для предотвращения проблем в учебном процессе.

3 **Оценка образовательных программ:** использование данных позволяет школам оценивать эффективность образовательных программ, выявлять сильные и слабые стороны учебного процесса, а также определять области для улучшения.

4 **Улучшение управления учебным процессом:** *Big Data* помогает администрации и педагогам принимать обоснованные решения на основе данных, улучшая планирование, ресурсное обеспечение и организацию образовательного процесса.

5 **Мониторинг прогресса учащихся:** данные могут использоваться для непрерывного мониторинга учебного прогресса учащихся, выявления областей, где необходима дополнительная поддержка, и своевременного реагирования на изменения.

Хотя использование *Big Data* в школах требует строгого соблюдения приватности данных и этических норм, правильное применение аналитики данных может значительно улучшить качество образования и помочь каждому учащемуся раскрыть свой потенциал.

Как улучшить конфиденциальность использования *BIG DATA* в школах. Улучшить конфиденциальность *BIG DATA* в школах можно путем применения следующих мер:

1 Обучение персонала: обучение администраторов, учителей и другого персонала школы по вопросам защиты данных и конфиденциальности поможет предотвратить случаи утечки информации.

2 Установка строгих прав доступа: регулирование уровней доступа к данным согласно принятой политики безопасности поможет предотвратить несанкционированный доступ к конфиденциальным данным учащихся.

3 Шифрование данных: шифрование хранимых данных поможет защитить информацию от несанкционированного доступа.

4 Анонимизация данных: при обработке и анализе данных следует использовать методы анонимизации, чтобы исключить возможность идентификации конкретных учеников.

5 Соответствие законодательству: соблюдение законодательных норм и требований по защите данных, таких как *GDPR*, поможет обеспечить законное использование информации об учащихся.

6 Мониторинг доступа к данным: регулярный мониторинг доступа к данным и аудит системы безопасности позволит оперативно выявлять и предотвращать нарушения конфиденциальности.

7 Проведение проверок на утечку данных: регулярное сканирование систем на наличие потенциальных утечек данных поможет обнаружить и устранить уязвимости в безопасности.

Методы использования *Big Data* в школе. Школа может собирать данные о предпочтениях учащихся, их успеваемости, стиле обучения, уровне знаний и предпочтениях в учебных предметах. Эти данные могут затем использоваться для:

1 Персонализации расписания занятий: на основе данных об учащихся можно разрабатывать индивидуализированные расписания занятий, учитывающие предпочтения каждого учащегося, их сильные и слабые стороны, что может способствовать повышению их мотивации и успеваемости.

2 Оптимизации нагрузки учащихся: анализ данных позволит более равномерно распределить учебные нагрузки на учащихся, учитывая их индивидуальные потребности и возможности.

3 Идентификации проблемных областей: алгоритмы анализа данных могут выявлять области, в которых учащиеся испытывают трудности, что поможет учителям адаптировать учебные материалы и методики обучения для более эффективного обучения.

4 Прогнозирование успеха: путем анализа данных о предыдущей успеваемости учащихся и их учебных показателях, можно прогнозировать их вероятность успешного завершения курса, что может помочь в своевременной поддержке и корректировке учебного процесса.

Данные методы использования *Big Data* в школе позволят персонализировать образовательный опыт для каждого учащегося, сделать обучение более эффективным и адаптированным под индивидуальные потребности, что в свою очередь может привести к повышению качества образования и успеваемости учащихся.

Использование помощи искусственного интеллекта в сборе информации *BIG DATA*. Использование нейронных сетей для сбора информации в образовательном контексте и *Big Data* может быть реализовано с помощью следующих методов:

1 Обработка больших объемов данных: нейронные сети могут быть задействованы для обработки и анализа больших объемов данных, например, результатов тестирований, анкет, учебных планов и записей учащихся.

2 Автоматизация процессов: использование нейронных сетей для автоматического сбора информации из различных источников, таких как онлайн-ресурсы, электронные журналы, отчеты учителей для выработки рекомендации по обновлению учебных планов и методик обучения.

3 Анализ поведения учащихся: нейронные сети могут помочь анализировать поведенческие данные учеников, такие как время занятий, ответы на задания, взаимодействие с учебными материалами, что позволит идентифицировать индивидуальные потребности каждого учащегося.

4 Предиктивная аналитика: нейронные сети могут быть использованы для создания моделей предиктивной аналитики, предсказывающих успеваемость учащихся на основе данных об их учебной деятельности и характеристик.

5 Персонализированное обучение: на основе собранных данных нейронные сети могут предложить персонализированные образовательные пути для учащихся, учитывая их индивидуальные особенности и потребности.

Использование нейронных сетей для сбора информации в образовательных учреждениях в сочетании с *Big Data* может значительно улучшить качество обучения, индивидуализировать процесс обучения и повысить успеваемость учащихся.

Заключение. Использование *Big Data* в образовании открывает возможности для инноваций и улучшения процесса обучения. Персонализация учебного процесса с помощью анализа данных позволяет адаптировать образовательную программу к потребностям каждого учащегося, делая обучение более эффективным и интересным. Прогнозирование успеваемости с помощью данных позволяет выявлять проблемные зоны и предоставлять своевременную поддержку учащимся.

Использование *Big Data* и персонализированных методов обучения не только повышает успеваемость учащихся, но и способствует развитию индивидуальных способностей каждого ребенка. Учебные заведения, которые активно внедряют анализ данных в образовательный процесс, открывают перед собой дверь к более эффективному и инновационному обучению.

Использование *Big Data* в образовании помогает создавать пространство для индивидуального роста каждого учащегося, обеспечивая более качественное и целенаправленное обучение в школах.

Список литературы

[1] Росляков А.В., Ваяшин С.В., Гребешков А.Ю. Интернет вещей: учебное пособие. Самара: ПГУТИ, 2015. – 200 с.

[2] Андреас Вайгенд. BIG DATA. Вся технология в одной книге.

[3] Оборудование инструментального производства, учеб. пособие / М. И. Михайлов, В. П. Кириленко ; М-во образования Респ. Беларусь, Гомел. гос. техн. ун-т им. П. О. Сухого. – Гомель : ГГТУ им. П. О. Сухого, 2016. – 270 с.

Авторский вклад

Дашкевич Марта Витальевна – руководство исследованием по использованию Big Data в образовании.

Дудко Диана Валерьевна – постановка задачи исследования, анализ существующих методов сбора и обработки информации в школе, изучение методов работы с BIG DATA, исследование конфиденциальности использования BIG DATA и применения искусственного интеллекта в сборе информации в школах.

BIG DATA IN EDUCATION: HOW THE USE OF DATA HELPS TO PERSONALIZE THE LEARNING PROCESS AND IMPROVE SCHOOL PERFORMANCE.

D.V. Dudko

Student of the State Educational Establishment “Lida RCTT” of the association of interests «Technology in Economics» and «Python Programming», student of the 10th grade of “Secondary School No. 16 named after P. M. Masherov in the city of Lida”

M. V. Dashkevich

Graduated from the Yanka Kupala State University of Grodno. The area of scientific interests is a project as one of the forms of research activities of students in the classes of interest associations.

Abstract. This article explores the role of Big Data in modern education with a focus on schooling. The ways of using data to personalize the educational process and improve student academic performance are analyzed. The methods of collecting and analyzing information about the preferences, needs and knowledge level of students to create individualized curricula are discussed. Examples of using Big Data to optimize class schedules, identify problem areas, predict student success and increase their motivation are presented. Practical examples of the use of data analytics in school education in order to improve the quality of education and individualize the approach to each student are considered. The main conclusion is that the use of Big Data makes it possible to effectively adapt the educational process to the needs of each student, which contributes to improving academic performance and the quality of education at school.

Keywords: Big Data, education, learning process, personalization, academic performance, school, confidentiality, data analysis, individualization of learning, optimization of schedules, motivation of students.

УДК 621.791.16

МОДЕЛИРОВАНИЕ ПРОВОЛОЧНОГО МОНТАЖА В МНОГОКРИСТАЛЬНЫХ МОДУЛЯХ



И.Б. Петухов

Начальник научно-технического центра
ОАО «Планар-СО»,
кандидат технических наук,
petuchov@kbtem.by



В.Л. Ланин

Профессор кафедры электронной техники и
технологии БГУИР,
доктор технических наук
vlanin@bsuir.by

И.Б. Петухов

Окончил Белорусский государственный университет по специальности радиофизика. Область научных интересов связана с исследованием проблем сборки изделий электронной техники. Начальник научно-технического центра ОАО «Планар-СО». Автор 4-х монографий, имеет 40 летний опыт работы в области сборки изделий электронной техники.

В.Л. Ланин

Окончил Минский радиотехнический институт. Профессор кафедры электронной техники и технологии. Автор 10 монографий, имеет 40 летний опыт работы в области технологии сборки и монтажа электронной техники.

Аннотация. Выполнен анализ факторов, определяющих стабильность термозвуковой микросварки золотой проволоки диаметром 17,5-50 мкм при сборке многокристалльных модулей. Определены зависимости влияния конфигурации сварочного капилляра на вид соединения и анализ алгоритма формообразования межсоединения.

Ключевые слова: проволочный микромонтаж, многокристалльные модули, ультразвуковая / термозвуковая микросварка, ультразвуковой преобразователь, сварочный капилляр.

Введение. Прогрессирующая миниатюризация многокристалльных модулей (МКМ) с обеспечением высокой функциональности вызвана, как техническими, так и экономическими соображениями. Очевидно, что набор определенных кристаллов, установленных на объединительной подложке (интерпозер) позволяет создать необходимый функционал в сжатые сроки даже без использования фотолитографии (рисунок 1). Топология интерпозера может быть выполнена установками лазерной обработки материалов с использованием режима лазерного выпаривания предварительно нанесенной металлизации с шириной линии снятия металлизации 15-20 мкм на кремниевой или керамической подложке.

Для обеспечения межсоединений кристалл–интерпозер–кристалл используется проволочный монтаж, который остается одним из основных методов сборки полупроводниковых приборов и интегральных микросхем (ИС), что объясняется высокой автоматизацией процесса, универсальностью по отношению к различным технологическим вариантам производства и геометрическими размерами изделий, а также более доступным технологическим процессом [1].

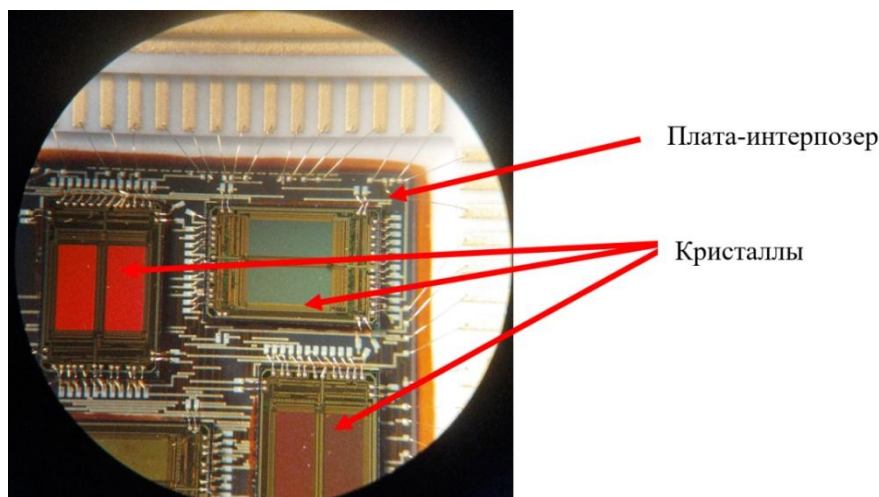
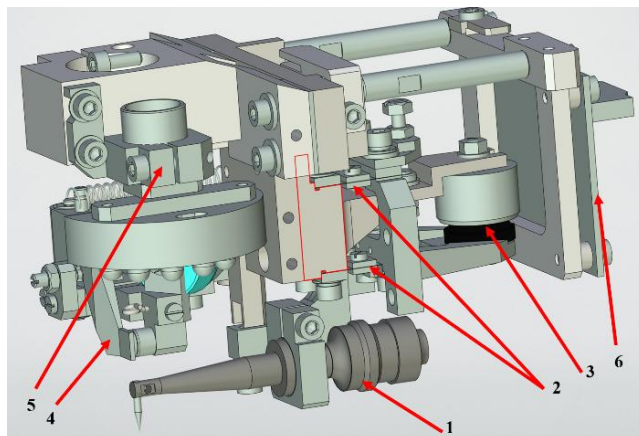


Рисунок 1. Фрагмент многокристальной сборки с использованием объединительной платы- интерпозера

В настоящее время монтаж золотых проволочных выводов диаметром от 17,5 до 50 мкм в основном осуществляется с помощью термозвуковой микросварки методами «клин-клин» и «шарик-клин» с подогревом рабочей зоны от 100 до 230 С, приложением заданного усилия к формируемому соединению проволока - контактная площадка и одновременным вводом ультразвуковых (УЗ) колебаний частотой в диапазоне 60-140 кГц через рабочий инструмент [2]. При монтаже золотой проволокой также используются методы термокомпрессионной (с температурой рабочей зоны в диапазоне 270-310 С) сварки и расщепленным электродом без воздействия УЗ колебаний при сборке структур, чувствительных к механическим вибрациям и появлению микротрещин в зоне образования соединения. На сегодняшний день практически на промышленном уровне освоена технология присоединения медной проволоки диаметром от 17,5 до 50 мкм методом термозвуковой сварки, но с необходимостью поддува защитного газа (95%N₂ 5%H₂) при формировании шарика на конце проволоки и в зоне соединения для защиты от окисления соединяемых материалов [3,4]. Существенным недостатком медной проволоки в данном случае является ее быстрая окисляемость, для предотвращения которой необходимо использование специальной тары хранения в азотной среде, а также короткие сроки использования проволоки. Для клиновой сварки широко используется алюминиевая проволока, которая также быстро окисляется, но свойства окислов алюминиевой проволоки и медной существенно различаются. Оксиды алюминия твердые и хрупкие и легко растрескиваются под действием ультразвуковых колебаний, в отличие от окислов на медной проволоке, которые мягкие и не разрушаются под действием ультразвука. Частным решением защиты медной проволоки является покрытие ее поверхности тонкой палладиевой пленкой.

Процесс термозвуковой микросварки золотой проволоки методом «шарик-клин» в технологическом плане обусловлен следующими параметрами: режимом формирования шарика на конце золотой проволоки, температурой рабочей зоны монтажа, амплитудой УЗ колебаний микросварочного инструмента (капилляра), контактным усилием (нагрузением) системы «торец капилляра – проволока - контактная площадка» и временем сварки. Очевидно, что чем тоньше используемая проволока, тем точнее должны задаваться исходные параметры и поддерживаться заданными в процессе присоединения. Применение тонкой проволоки < 20 мкм связано в первую очередь с размерами контактных площадок кристаллов < 60×60 мкм, а значит, и малой площадью соединения. Анализ технических характеристик установок шарикового и клинового присоединения

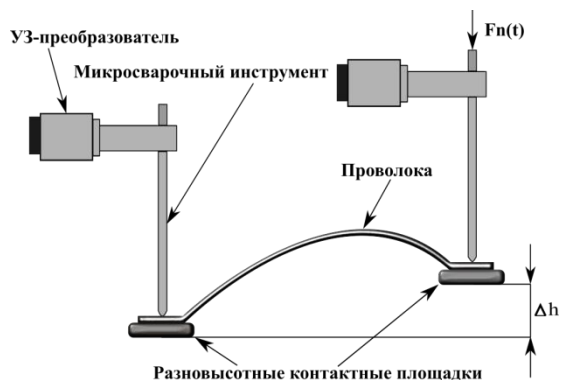
выводов ведущих мировых производителей, таких, как *Hesse & Knipps* [5], *Kulicke & Soffa* [6], *F&K Delvotec* [7] и *ASM* [8], показывает, что особое внимание при этом уделяется определению момента контактирования с точкой присоединения проволоки и управлению нагружением. Для 2,5D МКМ с заложенной в конструкцию прибора высокой разновысотностью Δh (до 4 мм) уровней сварки необходимо обеспечивать вертикальное положение микросварочного инструмента на позиции присоединения. Это может быть обеспечено перемещением сварочной головки (рисунок 2) по вертикальной оси Z.



1 – УЗ преобразователь; 2 – плоскопараллельная подвеска; 3- катушка нагружения, 4- механизм зажима проволоки, 5- держатель оптической системы, 6-кроссовое устройство

Рисунок 2. Сварочная головка для термозвуковой сварки золотой проволокой

Усилие $F_N(t)$ (рисунок 3) на микросварочный капилляр создается в большинстве случаев от электромагнитного актуатора - катушки в поле постоянного магнита по типу звуковой катушки (*voice coil motor*), ввиду простоты конструкции и возможности программирования посредством задания тока через катушку. Преимуществом подвески УЗ преобразователя на плоскопараллельных пластинах в сварочной головке является отсутствие трения качения и повышенная износостойкость по сравнению с узлами, выполненными на подшипниковых опорах, что делает их вариантом, привлекательным для применения в системах УЗ микросварки. Однако у подвески на плоскопараллельных пластинах есть и недостаток: зависимость усилия в зоне сварки от величины перебега сварочной головки (СВГ) после момента контактирования торца капилляра с точкой сварки. Перебег гарантирует необходимую степень деформирования проволоки для образования надежного соединения в процессе сварки и обычно его величина устанавливается равной диаметру проволоки. Фактически к предустановленному сварочному усилию прикладывается упругая сила изгиба плоскопараллельных пластин [9] и необходимо принимать меры по компенсации этого недостатка с помощью схемотехнических решений системы управления нагружением.



Δh - разновысотность контактных площадок

Рисунок 3. Положение УЗ преобразователя в сварочной головке на позиции присоединения

Микросварочный капилляр (инструмент) (рисунок 4) играет ключевую роль в термозвуковом присоединении золотой проволоки. Он обеспечивает точную и надежную сварку проволоки с контактными площадками прибора. Капилляр представляет собой иглу (рис.4б) со сквозным отверстием из керамики Al_2O_3 диаметром 1,59 мм и длиной L из ряда 9,5; 11,1; 12; 16 и 19 мм. Длина выбирается из необходимой глубины проникновения в изделие, например, если контактные площадки находятся в глубоком колодце. Стандартными считаются капилляры длиной 9,5 и 11,1 мм. Рабочий торец капилляра (рис.4а) определяет форму и размер сварного соединения типа «шарик» и «клин» и выбирается строго под заданный диаметр присоединяемой проволоки. Внешний угол рабочего торца FA для стандартных задач выбирается 4 или 8 градусов, угол внутренней фаски ICA 90 или 120 градусов.

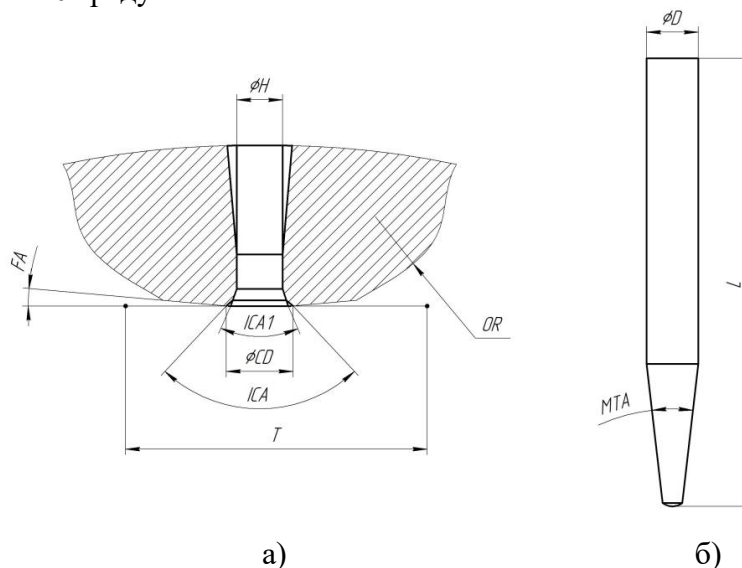


Рисунок 4. Рабочий торец микросварочного капилляра (а) и его общий вид (б)

Полученные экспериментальные данные для золотой проволоки диаметром 25 мкм отношения диаметра получаемых соединений к величине деформации шарика на контактной площадке показаны на рисунке 5.

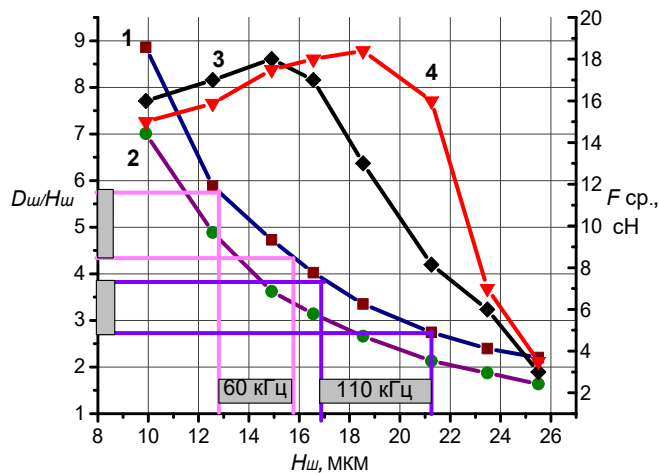


Рисунок 5.– Зависимости отношения диаметра деформированного шарика к его высоте (кривые 1 и 2) и величины прочности на сдвиг при ультразвуковых колебаниях 60 кГц (кривая 3) и 110 кГц (кривая 4) для проволоки диаметром 25 мкм

Для нестандартных задач, таких как, присоединение золотой, медной и алюминиевой проволоки методом «клин-клин» с использованием сварочной головки на рисунке 2, необходимо использовать угол рабочего торца FA 0 градусов, т.е. плоский торец, для обеспечения необходимой прочности и внешнего вида соединения. В этом случае (при монтаже проволоки по методу «клин-клин») режим формирования шарика на конце проволоки деактивируется и вместо этого формируется после второго соединения подогнутый под рабочий торец капилляра кончик проволоки, так называемый «хвостик», программируемой длины. Алгоритмы отрыва проволоки от места второй сварки (после формирования петли) и подгибки «хвостика» под торец могут быть разными, но важным моментом здесь является направление подгибки «хвостика», который всегда должен быть соосным с монтируемой петлей (проволочным стежком). Возможным и очевидным решением может быть введение возможности вращения сварочной головки вокруг своей оси.

Экспериментально установлено воздействие амплитуды ультразвуковых колебаний рабочего торца микросварочного капилляра на прямолинейность образуемых межсоединений, особенно при положении проволоки перпендикулярно направлению продольных ультразвуковых колебаний вдоль оси Y на рисунке 6. Поворот сварочной головки может значительно снизить этот эффект за счет поворота волновода УЗП вдоль оси образуемого соединения. Еще одним решением является использование двухуровневого ультразвукового импульса на второй точке присоединения вывода (рисунок 7). При этом длительность первого уровня T_1 составляет 10-15% от общей длительности T_2 , а амплитуда A_1 рассчитывается из следующего соотношения:

$$A_1 = A_2 \times \left(1 - \left| \frac{\cos \theta}{2} \right| \right) \quad (1)$$

где A_1 – амплитуда первого уровня ультразвукового импульса, A_2 – амплитуда второго уровня, θ – угол между направлением вывода и осью X (рисунок 6). Из соотношения (1) видно, что минимальная амплитуда A_1 достигается при $\theta = 0$, т.е. при перпендикулярном

направлении ультразвуковых колебаний относительно направления вывода, что снижает воздействие ультразвука на прямолинейность формируемого проволочного вывода.

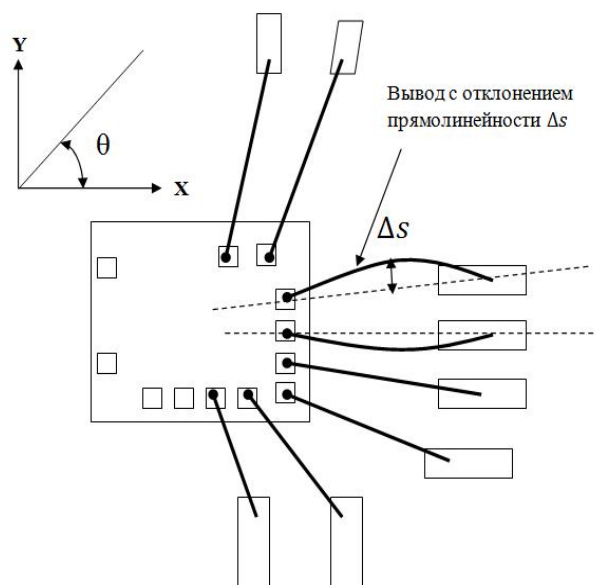


Рисунок 6. Отклонение вывода от прямолинейности на величину Δs при воздействии ультразвуковых колебаний торца капилляра вдоль оси Y

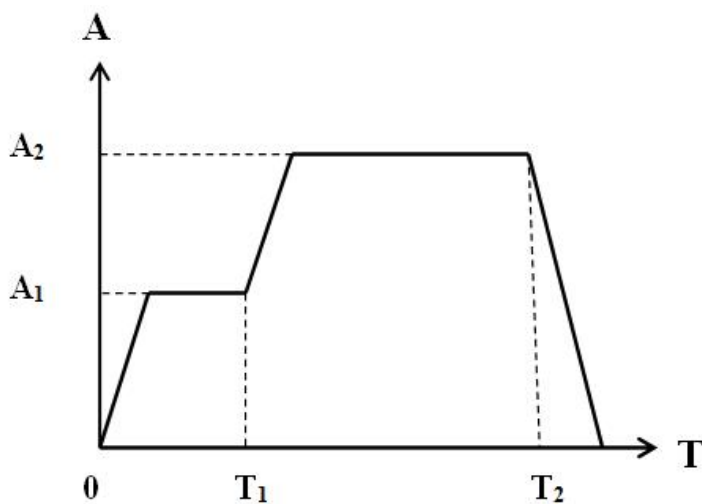


Рисунок 7. Профиль 2-х уровневый ультразвукового импульса на второй точке проволочного вывода, A - амплитуда УЗ колебаний; T - длительность

Выбор необходимого капилляра производится из каталога производителя, например, таких как *SPT* [10], *MicroPoint* [11], *GaiserTool* [12], которые являются лидерами в производстве микросварочных капилляров, инструментов монтажа кристаллов, расходных материалов и дополнительной оснастки. При выборе капилляра, особенно по его длине, надо иметь ввиду следующее. Капилляр зажимается в торец волновода УЗ преобразователя (УЗП) и необходимо быть уверенным после его установки в правильном согласовании с УЗП, поскольку капилляр является в данном случае нагрузкой УЗП.

Для оценки согласования системы УЗП – микросварочный капилляр на резонансной частоте придерживаются следующей методики. С помощью векторного анализатора импеданса снимают амплитудно-частотную характеристику (АЧХ) ультразвукового

преобразователя с поведением импеданса в районе резонансной частоты. Затем снимают АЧХ с зажатым капилляром в волноводе и добиваются минимального расхождения импеданса и резонансной частоты УЗП без капилляра и с установленным капилляром. Возможно при этом придется выбрать положение капилляра в волноводе (выше или ниже), так называемый вылет. Данную операцию рекомендуется выполнять с капиллярами других производителей, а также при проблемах стабильности сварочного процесса. Данная процедура не требует разборки сварочной головки и проводится достаточно оперативно. Типовые характеристики УЗП установки ЭМ-4520 в исполнении для термозвуковой сварки производства ОАО «Планар-СО» без инструмента и с установленным инструментом показаны на рисунке 8. Используемый капилляр имеет длину 11,1 мм, верхняя часть которого установлена заподлицо с волноводом. Из рисунка 6 видно, что изменение параметров ультразвукового преобразователя незначительное как по резонансной частоте ($F_s \sim 97,8$ кГц), импедансу ($R_l \sim 9$ Ом) и добротности ($Q_m \sim 430$), что говорит о хорошем согласовании системы ультразвуковой преобразователь-капилляр. Из практического опыта, существенными изменениями можно считать изменение частоты на 2 и более кГц, импеданса в 1,5 и более раз.

На приведенных диаграммах могут быть замечены побочные резонансы (в случае их наличия), которые могут создавать проблемы стабильной работе сварочной системе. Таким образом, данная процедура является оперативной диагностикой сварочной системы.

Формирование межсоединений в многокристалльных приборах достаточно сложная задача ввиду большого числа выводов и соответственно малым шагом между ними, значительной разновысотностью, соединениями различной длины и высоты.

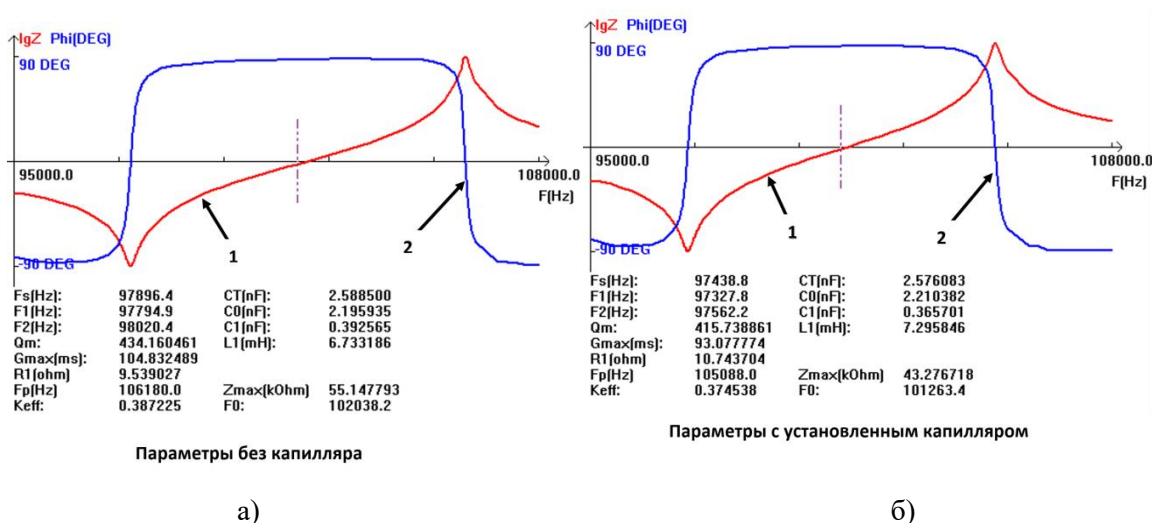


Рисунок 8. Параметры УЗ преобразователя без капилляра(а) и с установленным капилляром(б); 1-кривая импеданса; 2- кривая фазового сдвига между током и напряжением УЗП

Последовательность образования соединения контактной площадки кристалла и контактной площадки подложки изображена на рисунке 9. Из рисунка 9 видно, что для получения соединения необходимой формы капилляр совершает сложное интерполированное перемещение. На приведенном рисунке представлено последовательное перемещение капилляра для образования стандартного межсоединения треугольной формы. Катетами этой треугольной формы являются $(h+l_2)$ и l_1 . Начальное движение капилляра от первой точки соединения шариком встык производится в точку 1 и затем производится реверсивное движение в точку 2. Угол и высота перехода в точку 2

могут корректироваться в зависимости от физических параметров используемой проволоки.

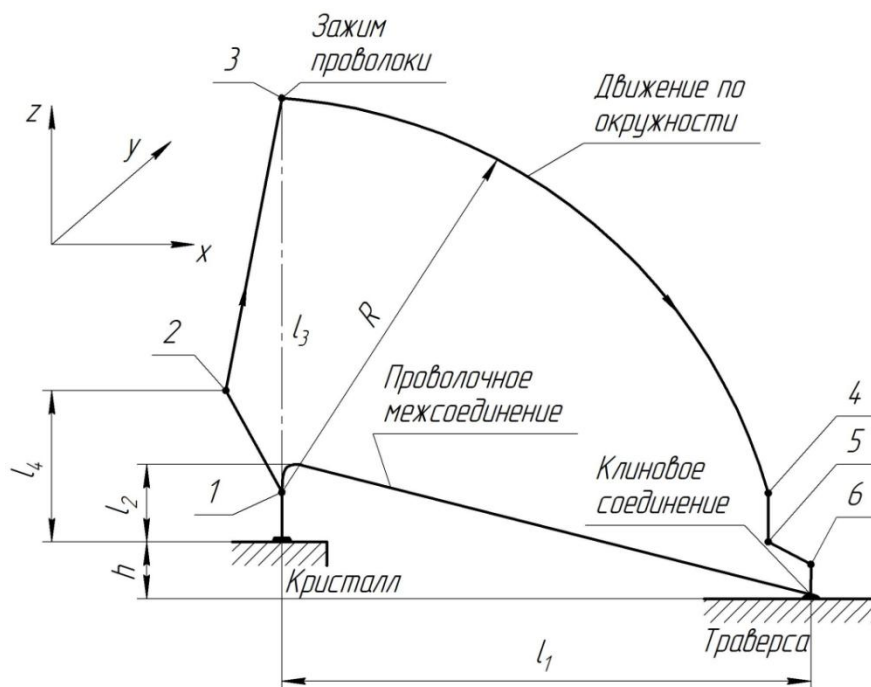


Рисунок 9. Последовательность перемещения капилляра при формировании проволочного соединения, где l_1 – длина межсоединения; l_2 – высота межсоединения; l_3 – высота от первой точки соединения до верхней точки подъема капилляра; l_4 – высота точки обратного реверса; R – радиус движения по окружности

Из точки 2 производится перемещение капилляра в точку 3 – верхнюю точку межсоединения. Здесь происходит по команде от управляющей системы зажим проволоки и перемещение по круговой траектории из точки 3 в точку 4. Далее следует вертикальное программируемое перемещение из точки 4 в точку 5. Заметим, что траектория 4-5 отстоит на некотором расстоянии от точки 6 – над местом присоединения второй точки сварки. Перемещением из точки 5 в точку 6 производится подтяжка петли для обеспечения ее прямолинейности. Очевидно, что все участки траектории должны быть рассчитаны с высокой точностью, а привод соответственно должен отработать эти перемещения, причем с высокой скоростью. Современные установки такое перемещение осуществляют за время 70-100 мс. Критичным параметром формирования межсоединения на рисунке 9 является верхняя точка подъема 3. Подъем капилляра ниже этой точки приводит к обрыву межсоединения, а выше – к искривлению формы. Из геометрии перемещения капилляра на рисунке 9 нетрудно определить, что высота подъема от места присоединения на кристалле в точку 3 определяется следующим образом:

$$l_3 = l_2 + (l_1 \times k) + \sqrt{(l_1^2 + (l_2 + h)^2)}, \quad (2)$$

где k – коэффициент в диапазоне 0...0,3.

Рассчитаем, например, высоту подъема l_3 (мм) для $l_2 = 0,3$ мм, $k=0$, $l_1 = 5$ мм и $h = 0,4$ мм. Подставляя значения в (2), получаем, что для длины межсоединения $l_1 = 5$ мм, высота подъема l_3 должна быть 5,35 мм. Коэффициент k вводится для корректировки формы петли для конкретного типа проволоки. Зависимости высоты подъема l_3 от длины межсоединения l_1 и разницы высот h уровней присоединения показаны на рисунке 10.

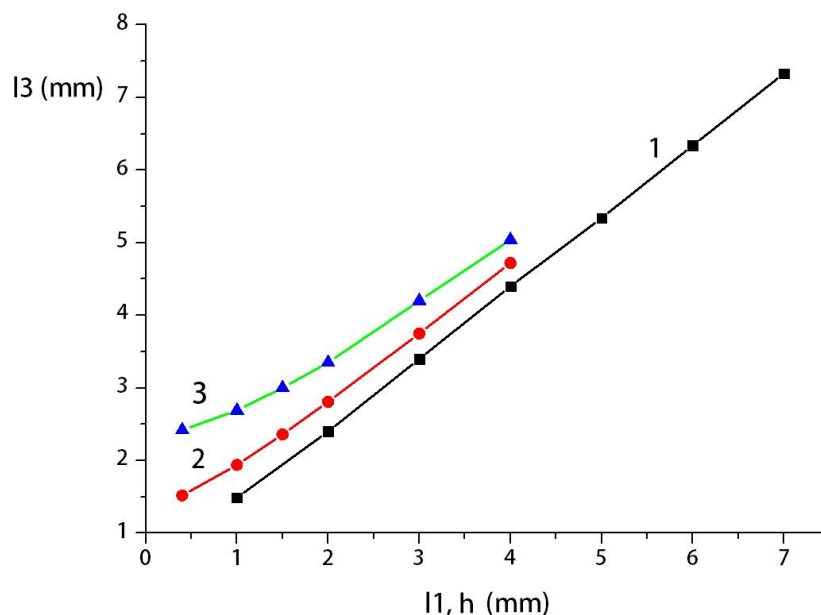


Рисунок 10. Зависимость высоты подъема l_3 от длины межсоединения l_1 (кривая 1), от разницы высот h для длины межсоединения 1 мм (кривая 2), от разницы высот h для длины межсоединения 2 мм (кривая 3)

Из рисунка 10 и формулы (2) легко определить исходные параметры перемещения сварочного капилляра, в частности высоты подъема для исключения обрыва проводника.

Выводы

1 Геометрия рабочего торца сварочного капилляра имеет ключевое значение при сборке многокристальных модулей. При использовании/замене капилляров от различных производителей рекомендуется проводить предварительное тестирование согласования с ультразвуковым преобразователем сварочной головки с использованием векторного анализатора, добиваясь минимального расхождения значений резонансной частоты и импеданса ультразвукового преобразователя с установленным капилляром и без него

2 Экспериментально установлено, что на прямолинейность образуемых межсоединений влияет направление ультразвуковых колебаний относительно оси проволоки, что требует корректировки амплитуды ультразвуковых колебаний сварочного капилляра за счет использования двухуровневого ультразвукового импульса на второй точке присоединения вывода, при этом длительность первого уровня составляет 10-15% от общей длительности, а амплитуда рассчитывается из следующего соотношения:

$$A_1 = A_2 \times \left(1 - \left|\frac{\cos \theta}{2}\right|\right),$$

где A_1 – амплитуда первого уровня ультразвукового импульса, A_2 – амплитуда второго уровня, θ – угол между направлением вывода и осью X.

Список литературы

- [1] Технологические процессы и системы в микроэлектронике: плазменные, электронно-ионно-лучевые, ультразвуковые / А.П. Достанко [и др.]. Минск: Бестпринт, 2009. –202 с.
- [2] Harmann, G.G. Wire bonding in microelectronics / G.G. Harman . Third Edition– N.Y. : McGraw Hill, 2010.– 427 p.
- [3] Cu wire and Pd-Cu wire package reliability and molding compounds / H. Abe [and etc.] // Proceeding 2012 IEEE 62nd Electronic Components and Technology Conference, IEEE, San Diego, CA. – pp. 1117-1123.
- [4] Fine pitch copper wire bonding in high volume production / В.К. Appelt [and etc.] // Microelectronics Reliability, Vol. 51 No. 1, 2011. - pp. 13-20.
- [5] Ultrasonic Wire Bonder for all wire sizes and applications! [Электронный ресурс]. – Режим доступа: <https://www.hesse-mechatronics.com/en/>
- [6] Innovative industry-leading ball bonding solutions [Электронный ресурс]. – Режим доступа: <https://www.kns.com/Products/Equipment/Ball-Bonder>
- [7] Range of products [Электронный ресурс]. – Режим доступа: <https://www.fkdelvotec.com/en/medien/downloads/>
- [8] We are ASMPT [Электронный ресурс]. – Режим доступа: <https://www.asmpacific.com/en/products?equipment=4>
- [9] Петухов, И.Б. Стабилизация сварочного усилия в процессе ультразвукового монтажа проволочных и ленточных выводов / И.Б. Петухов // Технологии и конструирование в электронной аппаратуре, 2021. – №1–2. –С. 49–53.
- [10] Semiconductor Tools – High Performance Ceramics HPC [Электронный ресурс]. – Режим доступа: <https://www.smallprecisiontools.com/en/homepage/?oid=55&lang=en>
- [11] Who we are [Электронный ресурс]. – Режим доступа: <https://mpptools.com/>
- [12] Gaiser catalog [Электронный ресурс]. – Режим доступа: <https://studylib.net/doc/25240729/gaiser-catalog>

Авторский вклад

Петухов Игорь Борисович – проведение экспериментальных работ, оптимизация алгоритмов формирования межсоединений при сборке многокристальных модулей.

Ланин Владимир Леонидович – постановка задачи исследования, анализ полученных результатов.

MODELLING WIRE BONDING IN MULTICHIP MODULES

I.B. Petuhov

Chief, R&D Centre of assembly equipment OJSC “Planar-SO”, PhD of Technical Sciences, Minsk Republic of Belarus, E-mail: petuchov@kbtem.by

V.L. Lanin

Professor, Department of Computer Design of BSUIR, Doctor of Technical Sciences

Abstract. The analysis of factors that determine the stability of thermosonic gold wire bonding with a diameter of 17.5-50 microns during the assembly of multichip modules was performed. Dependencies of the configuration of the bonding capillary on wire bonds and analysis of the algorithm of the loop formation are determined.

Key words: Wire bonding, multichip modules, ultrasonic / thermosonic bonding, ultrasonic transducer, bonding capillary.

УДК 621.383

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ БОЛЬШИХ ДАННЫХ НА ОСНОВЕ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ В СООТВЕТСТВИИ С ТРЕБОВАНИЯМИ ЗАКОНОДАТЕЛЬСТВА РБ



А.М. Тимофеев

Доцент кафедры защиты информации БГУИР,
кандидат технических наук,
доцент
TAMvks@mail.ru



А.Н. Шишпаренок

Учащийся учреждения образования «Национальный детский технопарк» по направлению «Информационная безопасность»
sashapistol22102008@gmail.com



В.Е. Юрут

Учащийся учреждения образования «Национальный детский технопарк» по направлению «Информационная безопасность»
vadimurut6@gmail.com

А.М. Тимофеев

Окончил Учреждение образования «Высший государственный колледж связи». Область научных интересов связана с фотоэлектронными процессами в фотоприемниках при одноквантовой регистрации, методами регистрации сверхслабых оптических потоков, а также комплексным обеспечением защиты информации в системах и сетях связи.

А.Н. Шишпаренок

Учащийся ГУО «Средняя школа №20 г. Борисова» (9 класс), учащийся учреждения образования «Национальный детский технопарк» в рамках индивидуальной учебной программы дополнительного образования одаренных детей и молодежи по направлению «Информационная безопасность» (22.11.2023 г. – 22.02.2024 гг.), научный руководитель доцент А.М. Тимофеев.

В.Е. Юрут

Учащийся ГУО «Средняя школа №3 г. Волковыска» (11 класс), учащийся учреждения образования «Национальный детский технопарк» в рамках индивидуальной учебной программы дополнительного образования одаренных детей и молодежи по направлению «Информационная безопасность» (22.11.2023 г. – 22.02.2024 гг.), научный руководитель доцент А.М. Тимофеев.

Аннотация. Применительно к системам обработки больших данных, содержащих персональные данные физических лиц, получена реализация метода введения идентификатора с учетом требований законодательства Республики Беларусь в сфере защиты информации.

Предложенная реализация позволяет выполнить обезличивание персональных данных, в результате чего нелегитимному пользователю системы становится невозможным без использования дополнительной информации определить принадлежность этих данных конкретному физическому лицу.

Ключевые слова: обезличивание, персональные данные, обработка персональных данных, субъект персональных данных, идентификатор.

Введение. В настоящее время существует потребность создания и обработки огромных массивов данных, которые содержат информацию различного назначения. Зачастую в составе таких массивов присутствуют персональные данные [1-4].

Под персональными данными будем понимать любую информацию, относящуюся к идентифицированному физическому лицу или физическому лицу, которое может быть идентифицировано [5].

Одной из основных задач при этом является обеспечение информационной безопасности персональных данных, поскольку собственник (владелец) информационной системы (или систем) обязан выполнять работы по защите информации этих данных, если они не являются общедоступными персональными данными [6].

Общедоступными персональными данными называют персональные данные, распространенные самим субъектом персональных данных либо с его согласия, либо распространенные в соответствии с требованиями законодательных актов [6].

Субъект персональных данных – это физическое лицо, в отношении которого осуществляется обработка персональных данных [6].

Обработка персональных данных – это действие или совокупность действий, совершаемых с персональными данными, включающие сбор, систематизацию, хранение, изменение, использование, обезличивание, блокирование, распространение, предоставление, удаление персональных данных [6].

Под обезличиванием будем понимать действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных к конкретному субъекту персональных данных [7, 8].

В соответствии с [6] для защиты персональных данных необходимо использовать методы их обезличивания. Одним из таких методов является метод введения идентификатора.

Метод введения идентификатора реализуется путем замены персональных данных, или части персональных данных, позволяющих идентифицировать субъект персональных данных, их идентификатором и создания таблицы соответствия с последующим раздельным хранением идентификаторов и таблиц [1]. В связи с тем, что до настоящего времени отсутствуют технические решения, которые позволяют выполнить обезличивание персональных данных, содержащихся в базах данных, относящихся к категории больших, целью данной работы являлась разработка технического решения по обеспечению информационной безопасности таких данных.

Объект исследования: структурированные и неструктурированные данные огромных объемов (большие данные).

Предмет исследования: применение метода введения идентификаторов для обеспечения информационной безопасности больших данных, содержащих персональные данные.

Реализация метода введения идентификатора. На рисунке 1 представлена программная реализация метода введения идентификатора.

При реализации метода введения идентификаторов важно, чтобы обезличенные данные можно было восстановить, применив процедуру деобезличивания [8].

Обезличенные данные – это действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных к конкретному субъекту персональных данных [8].


```
use bitvec::prelude::*;
use des::{Block, FromHexStr, MainKey, ToHexString};
use miette::{IntoDiagnostic, Result};
use std::io;

fn main() -> Result<()> {
    let main_key = MainKey::from_hex_str("AAAABBBBCCCCDDDD").into_diagnostic()?;
    let iv = Block::from_hex_str("FFFFFFFFFFFFFFFF").into_diagnostic()?;
    let iv = iv.encode(&main_key).into_diagnostic()?;

    println!("Input data in hex format (as one hex string): ");
    let input = get_input_string().into_diagnostic()?;
    let input = input.to_lowercase();
    let input = input.trim_start_matches("0x");
    let input = BitVec::from_hex_str(input).into_diagnostic()?;

    let output = encode(input, iv).to_hex_string();
    let (output, output_end) = output.split_at(output.len() - 16);
    let output_end = output_end.trim_start_matches('0');
    println!("\n{}", output, output_end);

    Ok(())
}

fn encode(input: BitVec, encoded_iv: Block) -> BitVec {
    let iv = encoded_iv.into_bitvec();
    let mut output: BitVec<usize, bitvec::order::LocalBits> = BitVec::with_capacity(input.len());
    let input = input.chunks(64);
    for chunk in input {
        let mut new_chunk = chunk.to_bitvec();
        new_chunk ^= iv.clone();
        output.extend(new_chunk);
    }
    output
}

fn get_input_string() -> io::Result<String> {
    let mut input = String::new();
    io::stdin().read_line(&mut input)?;
    Ok(input.trim().to_string())
}
```

исходные коды на языке программирования *Rust*

Рисунок 1. Реализация метода введения идентификаторов для обеспечения информационной безопасности больших данных, содержащих в своем составе персональные данные

Под деобезличиванием будем понимать действия, в результате которых обезличенные данные принимают вид, позволяющий определить их принадлежность к конкретному субъекту персональных данных, то есть, становятся персональными данными [8].

В этой связи применение методов, не обладающих свойством обратимости, нецелесообразно. Например, в случае реализации метода введения идентификатора с использованием односторонних функций хеширования (или односторонних функций, или дайджест-функций, или хеш-функций) отсутствует возможность выполнить процедуру деобезличивания [9-11].

В настоящей работе предложена и выполнена реализация метода введения идентификаторов, а также определены базовые принципы выполнения процедур обезличивания и деобезличивания персональных данных. Разработана компьютерная программа, которая позволяет выполнять процедуры обезличивания и деобезличивания персональных данных. В качестве языка программирования выбран язык *Rust* (рисунок 1). Отметим, что данная программная реализация может быть выполнена на любом языке программирования высокого уровня [12-17], поскольку не требует сложных математических операций.

Сущность технического решения, реализованного в компьютерной программе, заключается в следующем. Вначале персональные данные, подлежащие обезличиванию, кодируют с применением стандартных кодировочных таблиц, например, *UTF-16*. Важно отметить, что в качестве кодировочного формата может быть использована любая другая стандартная система, отличная от *UTF-16*, либо собственная («внутрифирменная») кодировочная таблица.

Затем формируют первый блок идентификатора ID_1 . Для этого вначале зашифровывают начальный вектор, который также называют вектором инициализации, или синхропосылкой [9, 10]. Далее применяют стандартный блочный шифр. В качестве такого шифра в работе использован стандарт *DES*, поскольку он прост в реализации, имеет достаточно высокий уровень информационной безопасности (с учетом схемных решений и режим его использования [9, 10]) и характеризуется высокой скоростью работы, что является весьма важным при обработке больших данных [1, 2]. Важно отметить, что вместо стандарта *DES* допустимо использовать стандарты *AES*, ГОСТ 28147-89, СТБ 34.101.31-2020 либо любой другой блочный шифр, имеющий криптостойкость, достаточную с учетом требований криптографической защиты информации, предъявляемых при эксплуатации информационной системы [9-11]. Результат зашифрования суммируют по модулю 2 с первым блоком закодированных персональных данных, получая первый блок идентификатора ID_1 .

Важно отметить, что разрядность блока закодированных персональных данных целесообразно выбирать равной разрядности двоичного числа, полученного на выходе используемого блочного шифра. Например, в случае применения СТБ 34.101.31-2020, разрядность такого блока может быть в диапазоне от 1 до 128 бит.

Аналогичным образом формируют остальные блоки идентификатора $ID_2 \div ID_N$ (где N – общее число блоков, определяемых объемом персональных данных), за исключением следующего. Перед зашифрованием очередного блока вначале обновляют входное значение, удалив старшие биты в количестве, равном выбранной разрядности блока. После этого оставшиеся биты сдвигают влево на такое число двоичных разрядов, а в освободившееся битовое пространство записывают предыдущий блок идентификатора. Например, при выработке второго блока ID_2 предыдущим блоком идентификатора является ID_1 .

Заключение. По результатам выполненной работы предложена реализация метода введения идентификаторов применительно к обезличиванию персональных данных, содержащихся в больших данных.

Определены принципы обеспечения информационной безопасности на основе криптографических и криптоподобных преобразований информации, полученная программная реализация метода введения идентификаторов для обеспечения информационной безопасности больших данных, содержащих в своем составе персональные данные.

Установлено, что полученные результаты могут быть использованы для обеспечения информационной безопасности персональных данных как при их обезличивании, так и

при деобезличивания, включая все основные этапы по обработке таких данных – сбор, систематизацию, хранение, изменение, использование и предоставление.

Работа выполнена при поддержке Учреждения образования «Национальный детский технопарк» (индивидуальная учебная программа дополнительного образования одаренных детей и молодежи по направлению «Информационная безопасность»).

Список литературы

- [1] Big Data = Большие данные : учеб. пособие / И. Б. Тесленко [и др.] ; Владим. гос. ун-т им. А. Г. и Н. Г. Столетовых. – Владимир: ВлГУ, 2021. – 123 с
- [2] Коллинз, М. Защита сетей. Подход на основе анализа данных / М. Коллинз. – Москва : ДМК Пресс, 2020. – 308 с.
- [3] Диогенес, Ю. Кибербезопасность: стратегии атак и обороны / Ю. Диогенес, Э. Озкаяя. – Москва : ДМК Пресс, 2020. – 326 с.
- [4] Семкин, С. Н. Основы правового обеспечения защиты информации : учеб. пособие для вузов / С. Н. Семкин, А. Н. Семкин. – М. : Горячая линия-Телеком, 2008. – 238 с.
- [5] Национальный реестр. Регистрационный номер 2/2819. Закон Республики Беларусь от 7 мая 2021 г. № 99-З «О защите персональных данных» [Электронный ресурс]. – Режим доступа : <https://pravo.by/document/?guid=12551&p0=H12100099>. – Дата доступа: 4.02.2024.
- [6] Национальный реестр. Регистрационный номер 7/4470. Приказ Оперативно-аналитического центра при Президенте Республики Беларусь от 20 февраля 2020 г. № 66 «О мерах по реализации Указа Президента Республики Беларусь от 9 декабря 2019 г. № 449» [Электронный ресурс]. – Режим доступа : <https://pravo.by/document/?guid=12551&p0=T62004470>. – Дата доступа: 4.02.2024.
- [7] Закон Республики Беларусь от 10 ноября 2008 г. № 455-З «Об информации, информатизации и защите информации» [Электронный ресурс]. – Режим доступа : [https://pravo.by/document/?guid=2012&oldDoc=2008-279/2008-279\(014-027\).pdf&oldDocPage=1](https://pravo.by/document/?guid=2012&oldDoc=2008-279/2008-279(014-027).pdf&oldDocPage=1). – Дата доступа: 4.02.2024.
- [8] Федеральная служба по надзору в сфере связи, информационных технологий и массовых коммуникаций Российской Федерации. Методические рекомендации по применению Приказа Роскомнадзора от 5 сентября 2013 г. № 996 «Об утверждении требований и методов по обезличиванию персональных данных» [Электронный ресурс]. – Режим доступа : https://rkn.gov.ru/docs/Xerox_Phaser_3200MFP_20131216122746.pdf. – Дата доступа: 4.02.2024.
- [9] Тимофеев, А.М. Криптографическая защита информации: учеб.-метод. пособие / А.М. Тимофеев. – Мн.: БГУИР, 2020. – 112 с.
- [10] Бутакова, Н.Г. Криптографические методы и средства защиты информации: учебное пособие / Н.Г. Бутакова, Н.В. Федоров. – СПб.: Интермедия, 2020. – 380 с.
- [11] Казарин, О.В. Надежность и безопасность программного обеспечения: учебное пособие / О.В. Казарин, И.Б. Шубинский. – М.: Юрайт, 2023. – 342 с.
- [12] Programming Rust / J. Blandy, J. Orendorff, L. Tindall. – Sebastopol, Ca: O'Reilly Media, 2021. – 1282 p.
- [13] Klabnik, S. The Rust Programming Language / S. Klabnik, C. Nichols. – San Francisco, Ca: No Starch Press, 2021. – 755 p.
- [14] Gjengset, J. Rust for rustaceans / J. Gjengset. – San Francisco, Ca: No Starch Press, 2022. – 283 p.
- [15] Прайс, М. C# 10 и .NET 6. Современная кросс-платформенная разработка / М. Прайс. – 6-е изд. – СПб.: Питер, 2023. – 848 с.
- [16] Керниган, Б. Практика программирования / Б. Керниган, Р. Пайк. – Москва ; Санкт-Петербург : Диалектика, 2019. – 288 с.
- [17] Плас, Дж. Вандер. Python для сложных задач : наука о данных и машинное обучение / Плас Дж. Вандер. – Санкт-Петербург : Питер, 2023. – 576 с.
- [18]

Авторский вклад

Тимофеев Александр Михайлович – научное руководство исследованием, постановка цели и задач исследования, разработка технического решения по реализации метода обезличивания персональных данных.

Шишпаренок Александр Николаевич – программная реализация метода обезличивания персональных данных, оценка обратимости процедуры обезличивания персональных данных (возможности проведения деобезличивания персональных данных), изучение принципов работы Big Data.

Юруть Вадим Евгеньевич – аналитический обзор литературных источников в рамках исследовательской работы, выполнение расчетной части, оценка свойств обезличенных данных, полученных по результатам предложенной в работе программной реализации метода обезличивания персональных данных, изучение принципов работы Big Data.

BIG DATA INFORMATION SECURITY

A.M. Timofeev

*Associate Professor of the
Department of Information
Security of BSUIR,
PhD of Technical Sciences,
Associate Professor*

A.N. Shyshparonak

*Student of the educational
institution «National Children's
Technopark» in «Information
Security»*

V.E. Yuruts

*Student of the educational
institution «National Children's
Technopark» in «Information
Security»*

Abstract. With regard to big data processing systems containing personal data of individuals, the implementation of the identifier input method has been obtained, taking into account the requirements of the legislation of the Republic of Belarus in the area of information protection.

The proposed implementation makes it possible to depersonalize personal data, as a result it becomes impossible for an illegitimate user of the system to determine the belonging of this data to a specific individual without the use of additional information.

Keywords: depersonalization, personal data, processing of personal data, personal data subject, identifier.

УДК 621.317.7+389.1 (075.8)

МОДЕЛИРОВАНИЕ ШИРОКОПОЛОСНОГО ФЕРРОМОДУЛЯЦИОННОГО МАГНИТОМЕТРА, ИНДИФФЕРЕНТНОГО К ХАОТИЧЕСКОМУ ВОЗБУЖДЕНИЮ ЧУВСТВИТЕЛЬНОГО ЭЛЕМЕНТА



В.Н. Ильин

Заведующий лабораторией
Института физики
им. Б.И. Степанова НАН
Беларуси, доктор технических
наук
ifanbel@ifanbel.bas-net.by



В.В. Баранов

Профессор кафедры
проектирования
информационно-
компьютерных систем
БГУИР, доктор технических
наук, профессор
vvb@bsuir.by



О.В. Дударева

Старший преподаватель
кафедры проектирования
информационно-
компьютерных систем
БГУИР, магистр наук,
kalita@bsuir.by

В.Н. Ильин

Окончил Минский радиотехнический институт (Белорусский государственный университет информатики и радиоэлектроники). Область научных интересов связана с исследованием и разработкой оптических систем контроля микродвижений технических объектов, моделированием систем скрытой передачи и приема информации.

В.В. Баранов

Окончил Минский радиотехнический институт (Белорусский государственный университет информатики и радиоэлектроники). Область научных интересов связана с моделированием и разработкой методов формирования твердотельных структур полупроводниковых приборов и интегральных микросхем, их тестированием, исследованиями в области био-медицинских технологий.

О.В. Дударева

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с моделированием твердотельных структур на основе гетеропереходов, исследованиями в области технических средств безопасности.

Аннотация. Проведен анализ возможностей реализации и моделирование широкополосного ферромодуляционного магнитометра, индифферентного к хаотическому возбуждению чувствительного элемента. Предложена схема модулятора сигналов, передаваемых в канал передачи информации.

Для проведения расчетов использована программа *MATHECAD*, в результате Фурье анализа напряжения на контуре получена спектрограмма сигнала модулятора. Показано, что прибор на основе предложенной схемы модулятора характеризуется линейной передаточной характеристикой в широком динамическом диапазоне, широкополосностью и отсутствием фазовых искажений.

Ключевые слова: Ферромодуляционный магнитометр, устойчивость к динамическому хаосу, схема и модель модулятора, феррозонд, факторы нестабильности.

Введение. Системы линейных дифференциальных уравнений, являющиеся приближениями по своей сути нелинейного Мира, благодаря своим упрощенным предположениям смогли описать многочисленные приложения и легли в основу их

моделей. Они были некогда единственно приемлемыми, так как только они отвечали духу времени, требованиям получения решения любым способом, но в замкнутом виде (в виде формулы). В связи с возможностями вычислительной техники при обработке больших объемов информации (*BIG DATA*), необходимость в упрощении моделей приобретает второстепенное значение и на первый план выдвигается адекватность математической модели сущности описываемого процесса. Стало возможным применять модели в виде систем нелинейных дифференциальных уравнений. И хотя их решения, как правило, не представимы в замкнутом виде, используя компьютер можно наблюдать за характером их поведения. При некоторых сочетаниях параметров модели ведут себя не предсказуемо. Решения в виде временных реализаций напоминают шум, их характер зависит от начальных условий. Такое явление получило название детерминированного или динамического хаоса. Оно часто встречается, но мало изучено, характерно для нелинейных колебательных систем и контуров. В измерительных технологиях детерминированный хаос может представить интерес в связи с применением активных датчиков (первичных преобразователей) в виде нелинейных колебательных контуров, активированных источниками возбуждения. Возникает вопрос, что является исключением в контуре, с параллельно включенными элементами $L(i)$, R , C , возбужденным источником тока, упрощенные процессы, или динамический хаос. На этот вопрос нельзя ответить потому, что хаотических режимов в контуре, бесконечно много, о чем свидетельствуют [1,2,3] численные и натурные эксперименты. Применение параметрического подхода, разработанного скорее для пояснения, чем для анализа принципов работы преобразователей, работающих в сугубо нелинейном режиме, является вопросом и пытаться получить ответ на этот вопрос – бессмысленно, так как он априори предполагает периодическое изменение параметра, а значит отсутствие хаоса. Таким образом, уверенность в том, что параметр $L(i)$ – (индуктивность, зависящая от тока, протекающего по катушке индуктивности) изменяется нужным образом и периодически, может быть обоснована, только в результате практических испытаний, но и здесь приходится сомневаться в том, не работает ли датчик на грани хаоса, оптимальным ли выбран режим. Определению доверительных интервалов изменения параметров датчика, может поспособствовать лишь его адекватная математическая модель. Как известно, шум ограничивает точность измерений, а появление хаоса может привести к катастрофе (невозможности проведения измерений и предсказания).

Безусловно, применительно к датчику, рассматриваемому ниже в качестве примера (магнитомодуляционному преобразователю), испытанному временем с характерными для подобного рода датчиков устоявшимися метрологическими показателями, многие из разработчиков которые не имели в свое время ни малейшего представления о детерминированном хаосе, проблемы связанные с его появлением, решались просто – подбором параметров и не представляли никакого интереса. Но дело не в конкретном, отработанном на сегодняшний день датчике, а о проблеме в целом, которая ранее ускользала от внимания и не замечалась. Приводимый ниже анализ, хотя и является частным решением задачи устойчивости системы к динамическому хаосу, однако его можно обобщить, учитывая, что нелинейных систем и явлений значительно больше, чем моделей их описывающих. Под обеспечением устойчивости в рассматриваемом случае понимаются не меры по предотвращению возможности возникновения хаоса, а работоспособная конструкция, схема, предполагающая его возникновение или существование.

Таким образом, целью доклада является доказательство принципиальной осуществимости измерения напряженности магнитного поля магнитомодуляционным преобразователем, характеризующимся динамическим хаосом, и разработка его модели индифферентной ко всем возникающим и существующим в нем режимам.

Модель и расчеты. Для измерения напряженности низкочастотных магнитных полей широко применяется феррозонд, представляющий собой прецизионный магнитный модулятор. Получили распространение его два вида, основанные на выделении второй гармоники частоты возбуждения и на измерении временных интервалов перемагничивания нелинейного элемента (ферромагнитного сердечника), по которым и судят о величине напряженности поля. Причиной, осложняющей понимания работы феррозонда, являются неизбежно присутствующие емкости, шунтирующие съемную обмотку и обмотку накачки. Емкости могут вводиться и специально. Например, в феррозонде на второй гармонике, для получения параметрического усиления. Как отмечается в [4, с. 65], «применять его все же следует с осторожностью, т.е. имели место случаи, когда феррорезонансная цепь возбуждения становилась источником появления ложного сигнала. Поэтому выбор феррорезонансного режима требует последующего тщательного выявления и устранения подобных негативных явлений». Приведенные слова разработчика подтверждают вышесказанное и приводят к необходимости рассмотрения уточненной модели феррозонда, т.е. в виде приближения нелинейным колебательным контуром. Для получения хаоса нелинейность необходима, но недостаточна. Необходимы еще и определенные сочетания параметров контура, частоты и амплитуды внешнего возбуждения. Что касается нелинейности, то следует отметить, что часто для рассмотрения феррозонда кривую перемагничивания приближают функцией $B = a \cdot H - b \cdot H^3$, где a и b – коэффициенты, B , H соответственно индукция и напряженность магнитного поля. [4]. На практике и графо - аналитическом рассмотрении пользуются кусочно - линейной аппроксимацией [4]. В этой связи для моделирования в *MATHCAD* можно предложить взаимно обратные линейно - кусочные функции, справедливые при любом значении аргумента [5].

$$H(B) = a \frac{B}{B_s} - 0.5 (-H_s + a) \left| \frac{B}{B_s} + 1 \right| - 0.5 (-a + H_s) \left| \frac{B}{B_s} - 1 \right|$$

$$B(H) = \frac{1}{a} \frac{H}{H_s} - 0.5 (-B_s + \frac{1}{a}) \left| \frac{H}{H_s} + 1 \right| - 0.5 (-\frac{1}{a} + B_s) \left| \frac{H}{H_s} - 1 \right|,$$

где B_s и H_s – соответственно индукция и поле насыщения сердечника, a – коэффициент,

$$b = \arctg\left(\frac{1}{a}\right), \quad \text{tg}(c) = \frac{H_s}{B_s}$$

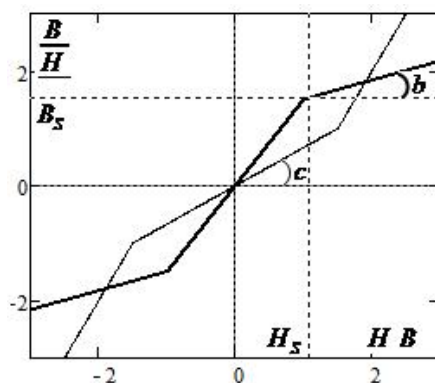


Рисунок 1. Зависимость B_s от H_s

Итак, рассмотрим магнитный модулятор в виде одного ферромагнитного сердечника, охваченного обмоткой, в которую подается ток возбуждения. Эта же обмотка будет и сигнальной.

Модулятор представлен на рис. 2 элементами $L(i)$, R , C , соединенными параллельно (блок-1). Параллельно этим элементам подключен генератор возбуждения 2 в виде источника тока $i(t)$ (H_B – напряженность внешнего измеряемого магнитного поля).

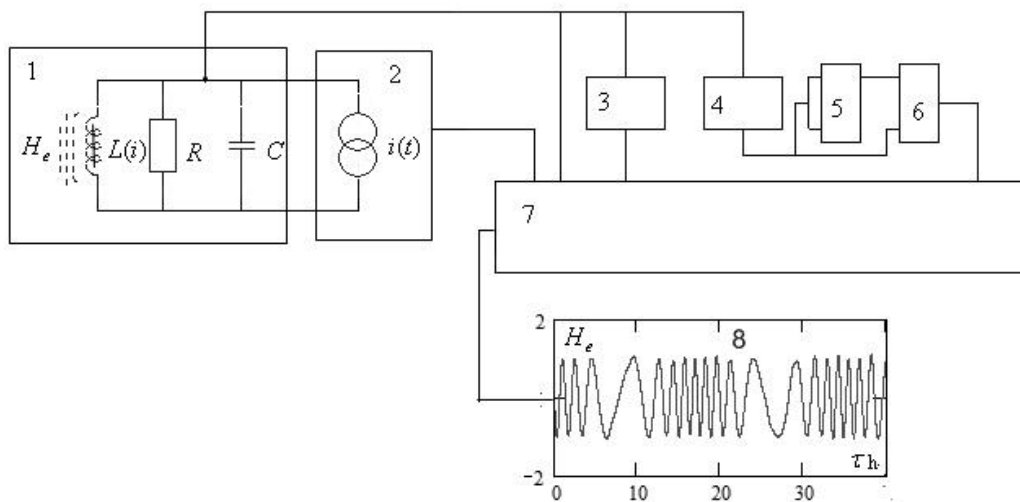


Рисунок 2. Схема магнитомодуляционного преобразователя:

- 1 – колебательный контур, 2 – генератор возбуждения,
- 3 – дифференциатор, 4 – интегратор, 5,6 – перемножители,
- 7 – сумматор, 8 – осциллограмма сигнала на выходе.

Для контура $L(i)$, R , C справедливо уравнение

$$C \frac{d^2 \phi}{dt^2} + \frac{1}{R} \frac{d\phi}{dt} + i_L = i, \quad (1)$$

где C – емкость контура, ϕ – магнитный поток в сердечнике катушки индуктивности $L(i)$,

i_L – ток в катушке индуктивности, R – эквивалентное сопротивление контура.

Введя обозначения:

$$\phi_0 = I_0 L_0; \frac{\phi}{\phi_0} = B; \omega_0^2 = \frac{1}{L_0 C}; \tau = \omega_0 t; p = \frac{\omega}{\omega_0}; K = \frac{1}{R} \sqrt{L_0 \frac{1}{C}} = \frac{1}{RC \omega_0}$$

$$i = I_m \sin(\omega t); H_m = \frac{I_m}{I_0}; H_L = I = \frac{i_L}{I_0},$$

где ϕ_0 – магнитный поток в сердечнике катушки индуктивности $L(i)$, i – величина тока в катушке и считая, что L_0 некая индуктивность, определяемая проходящим через нее током I_0 , уравнение контура запишем в безразмерном виде

$$\frac{d^2 B}{d\tau^2} + K \frac{dB}{d\tau} + I = H_m \sin(p\tau) \quad (2)$$

учитывая, что H_e напряженность внешнего измеряемого поля (безразмерная), присутствующая в катушке индуктивности, и отождествляя напряженность магнитного поля с током, создающим это поле ($H_L = I$), получим $H = H_L + H_B$.

Чтобы не отягощать рассмотрение и сделать его более прозрачным ограничимся, самой простой зависимостью $H = B^3$ [1,2,3,7]. Этого достаточно для возникновения хаоса, тогда (2) запишем окончательно в виде, удобном для компьютерного моделирования.

$$\frac{d^2 B}{d\tau^2} + K \frac{dB}{d\tau} + B^3 - H_e = H_m \sin(p\tau). \quad (3)$$

Следует отметить, что уравнение (3), в котором $H_e = 0$, привлекает исследователей своей обманчивой простотой, скрывая при этом глубокое внутреннее содержание, изучается на протяжении ряда лет, с целью установления критерия возникновения хаоса, однако сделано пока в этом направлении не очень много [1,2,3].

В натуральных и численных экспериментах хаос наблюдается в широком диапазоне значений параметров [1,2,3].

Создадим хаос, полагая, $H_e = 0.3$, $H_m = 5$, $p = 0.5$, $K = 0.005$,

$$h = \frac{m}{n} \cdot 2 \cdot \pi, n = 100000, m = 500, \tau = 1 \dots n$$

и решим уравнение (3) в программе *MATHCAD*.

Результаты решения представлены на рис.3. Совмещены две системы координат V , H и B , H_s , где V – напряжение на контуре. Эти фазовые портреты представлены точками отображения Пуанкаре, полученными в результате стробирования V , H , B с частотой второй гармоники генератора возбуждения (имитация синхронного детектирования применяемого в феррозондах). Индексы s при буквах на рисунке означают стробирование. В координатах V_s, H_s имеем бесформенное облако точек (хаос), В координатах B_s, H_s хаос прочерчивает инвариантную к нему кривую перемагничивания $B(H) = \sqrt[3]{H}$. Центр кривой перемагничивания смещен вправо на величину $H_e = 0.3$.

Таким образом, в данной ситуации, измерения, основанные на синхронном детектировании четных гармоник (второй), при действии на сердечник внешнего (измеряемого) магнитного поля ($H_e = 0.3$) становятся невозможными. То же можно сказать и о принципах измерений напряженности магнитного поля, основанных на расчетах временных интервалов перемагничивания сердечника, наблюдающихся в напряжении V [6].

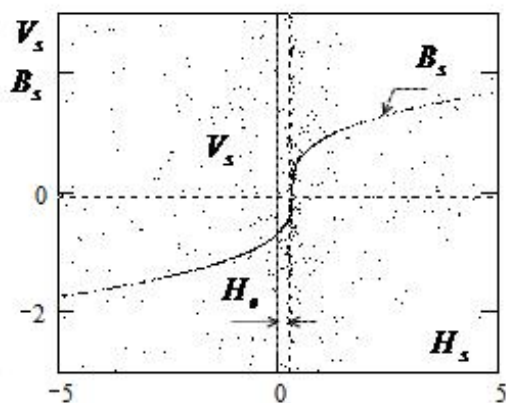


Рисунок 3. Графическое решение уравнения (3) в программе *MATHCAD*

Перейдем к рассмотрению возможной процедуры измерения. Отметим при этом, что магнитными модуляторами (впрочем, как и всеми модуляторами) измеряются магнитные поля в полосе частот несоизмеримо меньших частоты модуляции (возбуждения), поэтому для большего эффекта, убедительности и наглядности введем в контур сигнал с угловой модуляцией $H_e = 1 \cos(6 \tau + 6 \sin(0.7 \tau))$, полоса частот которого лежит в полосе частот детерминированного хаоса. Таким образом, верхняя граничная частота этого сигнала будет превышать частоту возбуждения модулятора приблизительно в 15 раз, что отражено на спектрограмме рис.4, полученной в результате Фурье анализа V при $m = 50$.

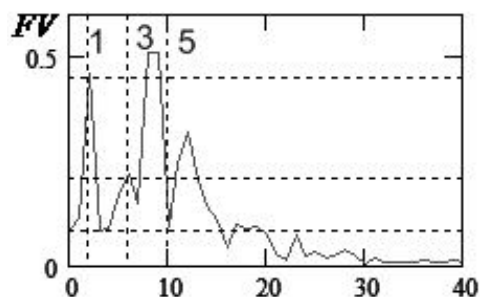


Рисунок 4. Спектрограмма сигнала модулятора

Сигнал H_e оказался полностью замаскирован хаосом. На хаотической спектрограмме FV (Фурье анализа V) (рис.4) видны экстремумы нечетных гармоник частоты возбуждения. Они обозначены цифрами. Цифра обозначает номер соответствующей гармоники. Согласно рис.4 частота первой гармоники на приведенной шкале частот равна 2.

Задача состоит в выделении из хаоса (измерении) сигнала H_e .

Как видно из уравнения (3), восстановить измеряемый сигнал H_e можно следующим образом.

Снять сигнал $V = \frac{dB}{d\tau}$ напряжения на контуре. Продифференцировать

V (R, C цепочкой), получить сигнал $\frac{d^2B}{d\tau^2}$ (емкостной ток). Это же V проинтегрировать (R, C цепочкой), получить B , проделав масштабирование V (умножая его на K), получить

ток IK через резистор. Далее индукцию B возвести в куб. Сделать это можно двумя перемножителями, с учетом, что каждый из них имеет по два входа. Затем эти сигналы сложить с сигналом тока накачки I_p , предварительно осуществив фазировку и масштабирование всех сигналов и получить информационный сигнал H_e .

В приводимом рассмотрении для нахождения напряженности магнитного поля в нелинейной индуктивности (возведения в куб $B(t)$) использовались перемножители. Известно, что при отсутствии гистерезиса зависимость $H(B)$ в общем случае может быть представлена и более сложной функцией в виде полинома нечетной степени $H = aB + cB^3 + dB^5 + \dots$, это не меняет сути, но усложняет схему, так как теперь придется увеличить число перемножителей, соединяя их определенным образом, а также применить сумматор имеющий большее число масштабируемых входов. Если наиболее подходящей для сердечника окажется кусочно-линейная аппроксимация, то вместо перемножителей можно воспользоваться функцией $H(B)$, обратной ограничителю $B(H)$, в принципе ее также можно реализовать на аналоговых элементах, диодах и операционных усилителях.

Вернемся к схеме магнитометра (рис.2). Сигнал адекватный току через емкость, получается в результате дифференцирования устройством 3 напряжения на контуре, продифференцированное напряжение подается на вход сумматора 7. На остальные входы сумматора приходят: напряжение с контура и напряжение, проинтегрированное блоком 4 после возведения его в куб перемножителями 5 и 6, а так же сигнал тока возбуждения с генератора 2. Сигналы масштабируются, складываются, и на выходе сумматора 7 на экране осциллографа 8 получаем измеряемый сигнал H_e . Можно отметить, что схема по своей структуре не сложнее традиционного магнитометра на второй гармонике с синхронным ее детектированием перемножителем, содержащая избирательный усилитель, настроенный на вторую гармонику частоты возбуждения.

Заключение. Показано, что в магнитном модуляторе, может наблюдаться детерминированный хаос. В этой связи следует отметить, что в процедуре известных методов модуляции, сигнал содержащий информацию должен быть отражен в изменении одного из параметров несущей (частоты, амплитуды или фазы пропорционально). Пока о каком-либо виде модуляции и подобии на него при динамическом хаосе говорить не приходится. Можно лишь констатировать о модуляции как о совместном результате сложного воздействия на модулятор информационного и гармонического сигнала накачки, и феноменологически описать принцип извлечения информационного сигнала из хаоса (осуществление демодуляции). Алгоритм действует, как в хаотическом, так и в конвергентном режимах работы модулятора.

Оказалось, что для создания магнитометра достаточно знать лишь кривую перемагничивания $H(B)$. Задача не является неожиданной, для нелинейной динамики она, так же как и выбор модели, является ключевой. Для ее решения требуются соответствующие исследования с помощью натурального эксперимента. Желаемая кривая перемагничивания может быть получена путем применения сердечников из специальных материалов с учетом их геометрических размеров. Можно проверить на модели (3), что нестабильность кривой перемагничивания в некоторых пределах допустима и не приводит к резкому изменению погрешности измерения.

На фоне современных успехов в магнитометрии, уделяющей прежде всего внимание разработке новых чувствительных к магнитным полям элементов, на основе различных физических эффектов, отмечать их недостатки и преимущества будет не целесообразно. Каждый прибор используется по своему назначению. Укажем лишь в этой связи, что потенциально, прибор на основе предлагаемого решения, может характеризоваться

линейной передаточной характеристикой в большом динамическом диапазоне, широкополосностью и отсутствием фазовых искажений. Следовательно, может применяться, как универсальная магнитная антенна, а также для регистрации импульсных магнитных полей с широким спектром частот, включая постоянную составляющую.

Подчеркнем, что в рассмотрении речь не идет о замене феррозондов предлагаемым решением, а прежде всего о принципиальной возможности их работы в режиме динамического хаоса. Альтернативу на сегодняшний день феррозонду на второй гармонике по чувствительности могут составить лишь квантовые магнитометры. Однако следует обратить внимание на то, что стоило только записать уравнение феррозонда в общем виде, отражающем его суть во всем известном спектре явлений, как концептуальная процедура измерения замаскированная в прошлом деформированной моделью, стала очевидной. Нелинейное уравнение феррозонда оказалось простым с виду, совсем не абстрактным, тем не менее, не поддающимся детальному разбору. Последнего удалось избежать, и обратить модель непосредственно в русло практического применения, не смотря на то, что детали и природа хаоса оказались так до конца и не выясненными.

Так как система (3) неинтегрируемая, и расчеты ее гармоник могут быть проведены приближенно разработанными для этих целей методами [5], то не исключено, что одновременно с регулярными наблюдаемыми траекториями, под ними находятся неразличимые, по результатам численного спектрального анализа Фурье, хаотические траектории. Гармоники не чистые, спектральные линии имеют всегда конечную ширину, существует дисперсия амплитуды, частоты и фазы гармоник. В таком случае (для этого в схеме не нужны никакие источники шумов), подобные артефакты приведут к нестабильности и дрейфу нуля приборов, основанных на детектировании гармоник. Явление, имеющее место в феррозонде на второй гармонике, отмечалось и изучалось многими авторами и приписывалось не без оснований некогерентному поведению магнитных доменов ферромагнитного материала при его переманчивании периодическим магнитным полем, иначе говоря, магнитному шуму. Не исключено, что некорректная настройка феррозонда, может привести к появлению в нем динамического хаоса, имитирующего те же явления.

Список литературы

- [1] Неймарк Ю.И., Ланда П.С. Стохастические и хаотические колебания. – М.: Наука, 1987.
- [2] Афанасьев Ю.В. Феррозондовые приборы. – Л.: Энергоатомиздат, 1986.
- [3] Семенов Н. М., Яковлев Н. И. Цифровые феррозондовые магнитометры. – Л.: Энергия, 1978.
- [4] Мун Ф. Хаотические колебания. – М.: Мир, 1990
- [5] Короткий В.П. Измерительный преобразователь в режиме динамического хаоса. – Измерительная техника, 2001, №11.
- [6] Хаяси Т. Нелинейные колебания в физических системах. – М.: Мир, 1968.
- [7] Горяченко В.Д. Элементы теории колебаний. – М.: Высшая школа, 2001.

Авторский вклад

Ильин Виктор Николаевич – руководство исследованием по разработке схемы и модели ферромодуляционного магнитометра, проведение расчетов в программе MATHCAD, построение по полученным результатам графиков.

Баранов Валентин Владимирович – постановка задачи исследования, описание работы модулятора, сравнение с другими опубликованными материалами, анализ полученных результатов расчетов.

Дударева Ольга Викторовна – участие в разработке модели модулятора, анализ литературных источников, оценка технических возможностей при практической реализации модулятора.

SIMULATION OF A BROADBAND FERROMODULATING MAGNETOMETER INDIFFERENT TO CHAOTIC EXCITATION OF A SENSING ELEMENT

V.N. Ilyin

*Head of the Laboratory
of Stepanov Institute of Physics
of the National Academy of
Sciences,
Doctor of Technical Sciences*

V.V. Baranov

*Professor of the Belarusian State
University of Informatics and
Radioelectronics, Doctor
of Technical Sciences, Professor*

O.V. Dudareva

*Assistant Professor of the
Belarusian State University of
Informatics and Radioelectronics,
Magister of Science*

Abstract. *An analysis of the possibilities of implementation and modeling of a broadband ferromodulation magnetometer, indifferent to the chaotic excitation of a sensing element, is carried out. A scheme of a modulator of signals transmitted to the information transmission channel is proposed. The MATHCAD program was used to perform the calculations, and as a result of the Fourier analysis of the voltage on the circuit, a spectrogram of the modulator signal was obtained. It is shown that the device based on the proposed modulator circuit is characterized by a linear transfer response in a wide dynamic range, wideband and the absence of phase distortions.*

Keywords. *Ferromodulation magnetometer, stable to dynamic chaos, modulator circuit and model, fluxgate, instability factors:*

УДК 658.5.012.7

ИСПОЛЬЗОВАНИЕ BIG DATA В ПРОИЗВОДСТВЕННОЙ АНАЛИТИКЕ



П.Р. Щербин

Студент кафедры «Технология машиностроения»

БНТУ

shcherbinpr@gmail.com

П.Р. Щербин

Студент кафедры «Технология машиностроения» машиностроительный факультет Белорусского национального технического университета. Область научных интересов связана с исследованием автоматизации производственных процессов.

Аннотация. Выполнен анализ платформ с низким кодом *Low-code*, *DATAMYTE*, которые позволяют быстро разрабатывать приложения для управления, анализа и визуализации больших данных производственных процессов. Показаны преимущества использования производственной аналитики для более эффективного принятия решений.

Ключевые слова: *Low-code*, *DATAMYTE*, *DataMetrics*, платформа с низким кодом, производственная аналитика.

Введение. Большие данные в обрабатывающей промышленности, основной столп 4.0, относятся к огромному объему данных, собранных из различных источников на протяжении производственного процесса, включая машины, продукцию и цепочки поставок. Эти данные, генерируемые с высокой скоростью и в различных форматах, подвергаются расширенной аналитике, машинному обучению и алгоритмам искусственного интеллекта.

Целью производства больших данных является получение ценной информации об эффективности, производительности и качестве, которая ранее была недоступна. Эти знания позволяют производителям принимать более быстрые и обоснованные решения, предвидеть потенциальные проблемы, оптимизировать процессы, а также стимулировать инновации и рост. Производители используют различные передовые технологии для сбора больших данных из различных источников на протяжении всей своей деятельности.

Датчики промышленного Интернета вещей (IIoT). Промышленный Интернет вещей – основополагающая технология для инициатив «Индустрия 4.0», которая использует подключенные интеллектуальные датчики, исполнительные механизмы и многое другое, объединяя персонал, продукты и процессы для обеспечения цифровой трансформации. Используя промышленные платформы Интернета вещей, компании подключают, отслеживают, анализируют промышленные данные и действуют на их основе новыми способами, чтобы повысить эффективность, максимизировать рост доходов, сократить расходы и многое другое. *IIoT* – это экосистема устройств, датчиков, приложений и сопутствующего сетевого оборудования, которые работают вместе для

сбора, мониторинга и анализа данных промышленных операций. Анализ таких данных помогает повысить наглядность и расширить возможности устранения неполадок и обслуживания. Это также может повысить эффективность, снизить затраты и повысить безопасность и защищенность [1].

Системы управления производством (MES). Системы управления производством – программные приложения, которые отслеживают и контролируют производственные процессы в заводских цехах. Они обеспечивают запись производственного процесса в режиме реального времени, собирая входные данные, результаты, состояние оборудования и данные о человеческом труде. Эти данные имеют неограниченное значение для выявления узких мест, оптимизации рабочих процессов и обеспечения постоянного улучшения.

Программное обеспечение для планирования ресурсов предприятия (ERP). ERP-системы – комплексные программные решения, которые интегрируют все аспекты бизнеса, включая производство, цепочку поставок, финансы и человеческие ресурсы. Они генерируют огромное количество данных в этих областях, предлагая целостное представление об операциях и облегчая принятие стратегических решений [2].

Промышленные системы управления. Системы промышленного контроля, включая системы диспетчерского управления и сбора данных и программируемые логические контроллеры, имеют решающее значение для управления автоматизированными процессами на производстве. Эти системы собирают данные от машин и производственных линий, контролируют их работу и обеспечивают их работу в пределах заданных параметров. Полученные данные можно анализировать для оптимизации эффективности производства и обслуживания оборудования. Хотя сбор больших данных в производстве является важным шагом, истинная ценность заключается в анализе этих данных.

Производственная аналитика тщательно изучает производственные данные, чтобы выявить идеи для более эффективного принятия решений. Для выявления закономерностей, прогнозирования и получения действенных бизнес-идей используются статистические модели, машинное обучение и искусственный интеллект.

Преимущества производственной аналитики [2]:

1 *Повышение эффективности.* Производственная аналитика помогает производителям выявлять неэффективность своей деятельности. Анализируя данные, собранные по оборудованию, сырью и рабочей силе, компании могут выявить области потерь или задержек и принять корректирующие меры. Это приводит к повышению эффективности производства и увеличению прибыли.

2 *Планирование профилактического обслуживания.* С помощью аналитики производители могут прогнозировать отказы оборудования еще до того, как они произойдут. Мониторинг и анализ данных датчиков *IIoT* позволяют компаниям выявлять аномалии, сигнализирующие о потенциальных проблемах, что позволяет им заранее планировать техническое обслуживание и предотвращать дорогостоящие простои.

3 *Оптимизация управления цепочками поставок.* Аналитика помогает оптимизировать операции цепочки поставок путем анализа данных о спросе, запасах и логистике. Это позволяет производителям эффективно балансировать спрос и предложение, оптимизировать управление запасами и минимизировать затраты.

4 *Улучшение качества продукции.* Анализируя данные с различных этапов производственного процесса, производственная аналитика может помочь обеспечить качество продукции. Это может помочь выявить факторы, способствующие возникновению дефектов, что позволит производителям усовершенствовать свои процессы и поставлять высококачественную продукцию.

5 *Инновации.* Информация, основанная на данных, может стимулировать инновации,

выявляя возможности для новых продуктов или улучшений существующих процессов. Эти идеи могут дать производителям решающее преимущество в высококонкурентной отрасли.

6 *Помощь в принятии стратегических решений.* Производственная аналитика обеспечивает эмпирическую основу для принятия стратегических решений. Производители могут делать осознанный выбор в отношении производственных графиков, распределения ресурсов и будущих инвестиций, анализируя тенденции, закономерности и корреляции в данных. Это способствует развитию культуры эффективного принятия решений, когда стратегии тщательно разрабатываются для достижения оптимальных результатов.

Оптимизация больших данных в производстве с помощью платформы Low-code.

В эпоху Индустрии 4.0 интеграция платформы с низким кодированием может изменить правила игры в оптимизации больших данных в производстве. Эти платформы упрощают создание и поддержку приложений для управления, анализа и визуализации больших данных, делая решения на основе данных более доступными, быстрыми и экономичными [2, 4].

Платформы с низким кодом позволяют быстро разрабатывать приложения, предоставляя удобный интерфейс с готовыми шаблонами, функциями перетаскивания и интуитивно понятными элементами дизайна. Это позволяет производителям быстро создавать и развертывать решения по управлению данными, адаптированные к их уникальным требованиям, без необходимости тщательного кодирования.

Упрощая процесс создания инструментов анализа данных, платформы с низким уровнем кода демократизируют анализ данных. Они позволяют нетехническим пользователям создавать, изменять и использовать приложения для анализа данных, позволяя более широкому персоналу использовать понимание больших данных в своей работе.

Платформы *Low-code* часто включают в себя комплексные инструменты визуализации данных, которые могут преобразовывать сложные наборы данных в простые для понимания диаграммы, графики и информационные панели. Это может помочь производителям получить целостное представление о своей деятельности и более эффективно принимать решения на основе данных [2, 4].

Платформы *Low-code* предназначены для поддержки масштабируемости. По мере роста бизнеса и увеличения объемов данных приложения, созданные на этих платформах, можно легко модифицировать и масштабировать в соответствии с новыми требованиями. Это означает, что они остаются пригодными для использования даже при изменении потребностей производственной организации в больших данных.

Благодаря встроенным функциям безопасности платформы *Low-code* помогают производителям гарантировать, что их приложения для управления данными соответствуют соответствующим стандартам и нормам безопасности данных. Автоматические обновления и исправления также означают, что эти приложения остаются защищенными от развивающихся угроз.

Интеграция низкокодовой платформы для управления большими данными в производстве может привести к трансформации. Это может помочь производителям более эффективно использовать весь потенциал своих данных, обеспечивая прочную основу для цифровой трансформации производства.

Оптимизация больших данных в производстве с помощью DATAMYTE.

DATAMYTE – платформа управления качеством с возможностью минимального использования кода. *DataMetrics* – это программное обеспечение для автоматизации рабочих процессов с низким уровнем кода, встроенным контрольным списком и интеллектуальным конструктором форм. Эта функция и ее интерфейс с функцией

перетаскивания позволяют создавать любые необходимые вам контрольные списки и шаблоны форм, в том числе специально предназначенные для оптимизации процессов обработки производственных данных.

DATAMYTE позволяет проводить многоуровневый аудит процессов, систематический анализ важнейших этапов процесса. Этот метод аудита фокусируется на областях с наибольшим риском сбой или несоблюдения требований. Проведение многоуровневого аудита процессов (*LPA*) с использованием *DATAMYTE*, позволяет эффективно выявлять и исправлять дефекты до того, как они усугубят производственный процесс [3].

Для сбора данных и статистического контроля процессов используется программное обеспечение *DataMetrics* (рисунок 1), которое поддерживает универсальный сбор данных посредством ручного, полуавтоматического, автоматизированного и портативного сбора данных. Статистический контроль процессов (*SPC*) – метод измерения и контроля качества путем мониторинга производственного процесса. Данные о качестве собираются в форме измерений продукта или процесса или показаний различных машин или приборов. Данные собираются и используются для оценки, мониторинга и контроля производственного процесса.

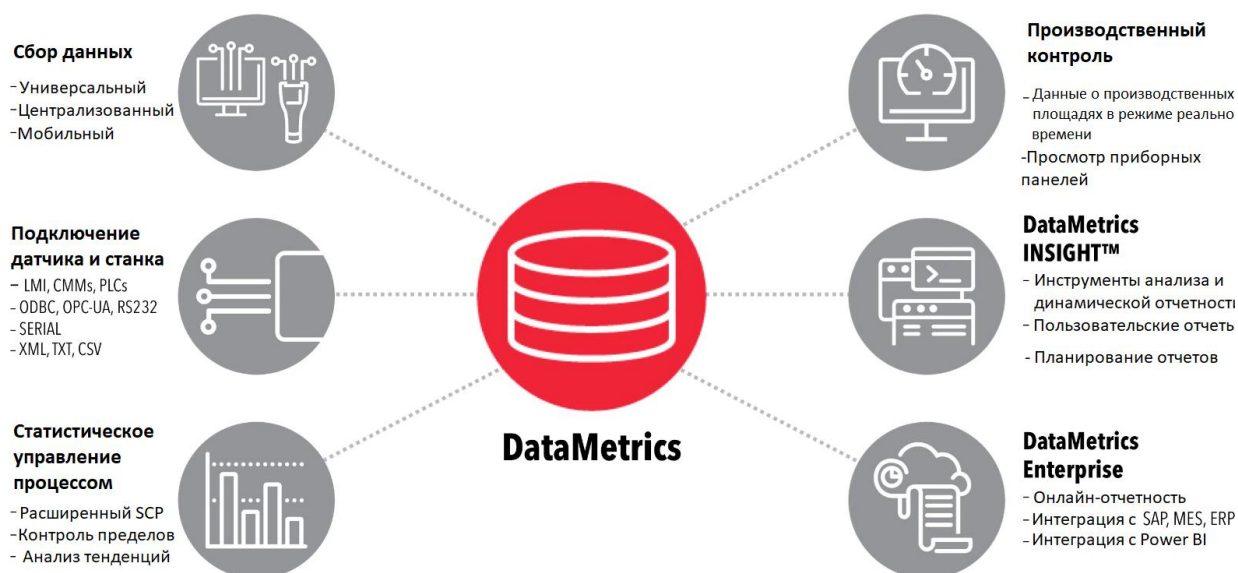


Рисунок 1. Система DataMetrics [3]

DataMetrics предоставляет возможность собирать все данные о качестве производственного процесса в одну центральную базу данных и составляет отчеты по этим данным. Анализ и отчетность, поддерживаемые базой данных, обеспечивают целостность данных и отслеживаемость процессов, что соответствует самым строгим стандартам во многих отраслях.

Простые в разработке веб-отчеты на информационной панели в режиме реального времени обеспечивают полную видимость процесса и оптимизацию производства [3].

Статистический контроль и сбор данных *DataMetrics* предоставляют возможность использовать упреждающий подход для поддержания или улучшения качества продукции, достижения производственных целей, сокращения затрат на брак и доработку или соблюдения OEM-стандартов или отраслевых стандартов.

Заключение. Большие данные в сочетании с возможностями *Low-code* платформ производят революцию в производственных операциях во всем мире. Эти достижения упрощают анализ данных, улучшают визуализацию и обеспечивают масштабируемость и безопасность. Освоение этой цифровой трансформации является ключевым моментом для производителей, стремящихся к беспрецедентной эффективности и контролю качества.

DATAMYTE – идеальное решение для создания и внедрения инструментов на основе данных для контроля качества, улучшения процессов, а также здоровья и безопасности сотрудников.

Список литературы

[1] CISCO [Электронный ресурс]. – Режим доступа: <https://www.cisco.com/c/en/us/solutions/internet-of-things/what-is-industrial-iiot.html>– Дата доступа: 11.02.2024.

[2] DATAMYTE [Электронный ресурс]. – Режим доступа: <https://datamyte.com/blog/big-data-in-manufacturing-industry/>– Дата доступа: 11.02.2024.

[3] DATAMYTE [Электронный ресурс]. – Режим доступа: <https://datamyte.com/datamyte-datametrics/>– Дата доступа: 11.02.2024.

[4] IBM [Электронный ресурс]. – Режим доступа: <https://www.ibm.com/topics/low-code>– Дата доступа: 11.02.2024.

USING BIG DATA IN PRODUCTION ANALYTICS

P.R. SHCHerbin

*Student of the Department of Mechanical
Engineering Technology at BNTU*

Abstract. The analysis of Low-code, DATAMYTE platforms has been performed, which allow you to quickly develop applications for managing, analyzing and visualizing big data of production processes. The advantages of using production analytics for more effective decision-making are shown.

Keywords: Low-code, DATAMYTE, DataMetrics, low-code platform, production analytics.

УДК 004.94:62-9-027.45

СТАТИСТИЧЕСКОЕ ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ В ИССЛЕДОВАНИИ ЭФФЕКТИВНОСТИ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ НАДЁЖНОСТИ ИЗДЕЛИЙ ПО ИНФОРМАТИВНЫМ ПАРАМЕТРАМ



С.М. Боровиков

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

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.

Аннотация. Рассматривается применение статистического имитационного моделирования для исследования эффективности моделей прогнозирования класса работоспособности изделий электронной техники для заданной наработки по результатам контроля информативных параметров изделий в начальный момент времени. Особенностью исследуемых моделей прогнозирования является преобразование информативных параметров в дискретный двоичный или троичный код. Рассматривается несколько способов построения моделей на основе использования двоичных или троичных сигналов. Модели получены с помощью обработки данных обучающего эксперимента большого объёма, смоделированного на компьютере, а эффективность моделей определена путём их применения как к экземплярам обучающей выборки, так и к экземплярам смоделированных контрольных выборок.

На основе исследований даны рекомендации по использованию моделей для решения практических задач прогнозирования класса работоспособности изделий для заданных наработок.

Ключевые слова: информативные параметры, преобразование в кодовые сигналы, класс надёжности изделий, модели прогнозирования.

Введение. Решение о надёжности технических изделий, в том числе изделий электронной техники (ИЭТ) при использовании информативных параметров, контролируемых в начальный момент времени, принимают в виде отнесения прогнозируемого изделия (конкретного экземпляра) к одному из двух классов с точки зрения работоспособности (надёжности) этого экземпляра для заданной наработки (времени работы) t_n :

– K_1 – класс работоспособных экземпляров (экземпляров повышенного уровня надёжности);

– K_2 – класс неработоспособных экземпляров (потенциально ненадёжных экземпляров).

Практика показала, что использование одного информативного параметра может привести к заметным ошибкам прогнозирования в виде «перепутывания» по результатам

индивидуального прогнозирования номера класса экземпляров, к которому они в действительности будут принадлежать на момент наработки t_n . Поэтому для прогнозирования используют несколько информативных параметров. Их число k в основном выбирают в диапазоне $k = 2 \dots 5$ [1].

Совокупность информативных параметров $x_1^{(j)}, x_2^{(j)}, \dots, x_k^{(j)}$, измеренная в начальный момент времени у конкретного экземпляра (будем называть его j -м экземпляром) создаёт образ этого экземпляра. Задача индивидуального прогнозирования в этом случае состоит в том, чтобы правильно распознать этот образ, т.е. верно указать класс (K_1 или K_2), к которому для интересующей наработки t_n будет принадлежать j -й экземпляр. Поэтому методы индивидуального прогнозирования надёжности ИЭТ по их информативным параметрам известны также как методы «распознавания образов».

Для выполнения индивидуального прогнозирования класса изделий для заданной наработки t_n нужно располагать моделью прогнозирования. Её получают, используя предварительные экспериментальные исследования на надёжность некоторой выборки изделий интересующего типа. Эти исследования называют обучающим экспериментом.

Модель прогнозирования обычно получают в виде [1, 2]

$$\begin{aligned} j & K_1, \text{ если } F[x_1^{(j)}, \dots, x_k^{(j)}] \geq P_0, \\ j & K_2, \text{ если } F[x_1^{(j)}, \dots, x_k^{(j)}] < P_0, \end{aligned} \quad (1)$$

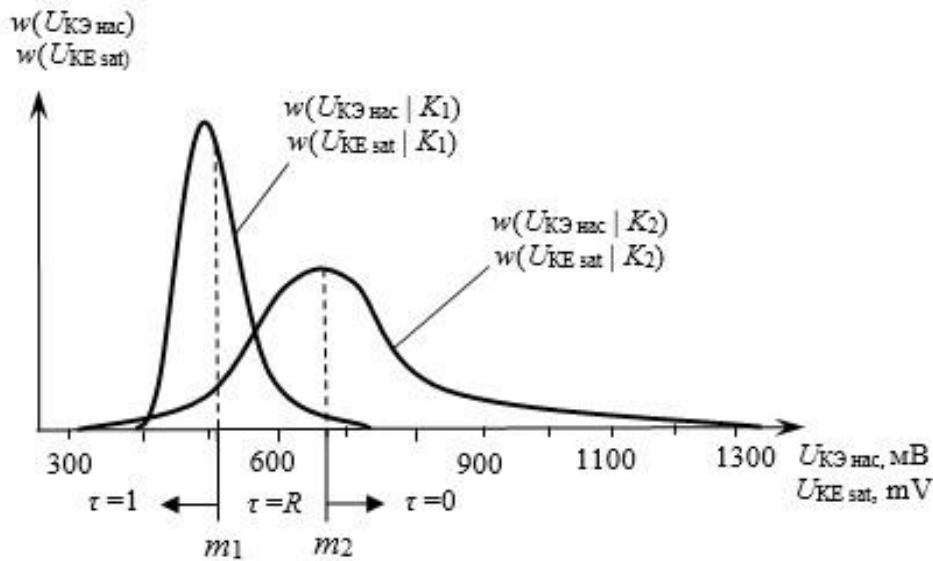
где j – символ, используемый для обозначения конкретного экземпляра; $x_1^{(j)}, \dots, x_k^{(j)}$ – значения информативных параметров, измеренные для j -го экземпляра в момент времени $t = 0$; $F[x_1^{(j)}, \dots, x_k^{(j)}]$ – значение прогнозирующей функции, подсчитанной для j -го экземпляра; k – число информативных параметров, используемых для решения задачи индивидуального прогнозирования класса экземпляров для заданной наработки; P_0 – порог (иначе – критерий) разделения классов, определяемый экспериментально.

Полученную модель вида (1) применяют к однотипным экземплярам, не принимавшим участия в обучающем эксперименте.

Известные методы прогнозирования [1, 2], такие как метод статистических решений, метод потенциальных функций и некоторые другие являются сложными как для понимания, так и для автоматизации процедуры прогнозирования.

Известно [2], что непрерывные отчёты информативных параметров ИЭТ в определённой степени содержат избыточное количество информации о классе работоспособности экземпляра для интересующей наработки t_n . С учётом этого представилось возможным перейти на двоичное представление информативных параметров. Тогда образ прогнозируемого экземпляра можно описать набором двоичных чисел и далее выполнять дальнейшую их логическую обработку. Используя принципы пороговой логики [3], модель прогнозирования можно получить в виде логической таблицы, показывающей какой комбинации двоичных сигналов соответствует тот или иной класс работоспособности ИЭТ [1, 4, 5].

Актуальность исследований. Установлено, что лучшие результаты прогнозирования обеспечивает преобразование информативных параметров в троичный код [1, 6, 7]. Такое преобразование иллюстрируется рисунком 1 на примере информативного параметра $U_{КЭнас}$ (напряжение насыщения коллектор эмиттер) биполярных транзисторов типа КТ872А.



m_1, m_2 – среднее значение (математическое ожидание) параметра соответственно для экземпляров классов K_1 и K_2

Рисунок 1. Преобразование информативного параметра $U_{КЭ нас}$ в троичный код

В случае троичного преобразования для каждого информативного параметра x_i рассматривается три области его изменения и каждой из них присваивается свой кодовый троичный сигнал τ_i (см. рисунок 1). Для области, находящейся между средними значениями информативного параметра в классах K_1 и K_2 , т.е. для области от m_1 до m_2 , присваиваем код $\tau_i = R$ (от англ. Range – диапазон). Для этой области характерна высокая степень неопределённости класса работоспособности рассматриваемого ИЭТ (экземпляра). За пределами диапазона от m_1 до m_2 кодовый сигнал $\tau_i = 1$ (единица) присваивается той области информативного параметра, значения которой в основном соответствуют экземплярам класса K_1 , а код $\tau_i = 0$ (нуль) – области информативного параметра, значения которой в основном соответствуют экземплярам класса K_2 . С учётом этого алгоритм преобразования информативных параметров в троичный код запишется в виде [6, 7]

$$\begin{aligned} \Phi &= 1, \text{ если } x_i > m_1 \\ \Phi &= 0, \text{ если } x_i < m_2 \\ \Phi &= R, \text{ если } m_2 < x_i < m_1. \end{aligned} \quad (2)$$

Соотношения (2) должны использоваться для информативных параметров в случае выполнения условия $m_1 > m_2$. Если имеет место $m_1 < m_2$, то следует использовать соотношения

$$\begin{aligned} \Phi &= 1, \text{ если } x_i < m_1 \\ \Phi &= 0, \text{ если } x_i > m_2 \\ \Phi &= R, \text{ если } m_1 < x_i < m_2. \end{aligned} \quad (3)$$

Общая модель прогнозирования в случае преобразования информативных параметров в троичные сигналы 1, 0 и R может быть записана в виде, аналогичном модели метода пороговой логики [1, 3, 4], согласно которому информативные параметры

преобразуются в двоичные сигналы 1 (единица) или 0 (нуль). При троичном преобразовании модель индивидуального прогнозирования примет вид

$$D^{(j)} = \begin{cases} 1, & \text{если } F^{(j)} = \mathbb{W}[\mathfrak{b}_i, \mathfrak{f}^{(j)}] \geq T, \\ 0, & \text{если } F^{(j)} = \mathbb{W}[\mathfrak{b}_i, \mathfrak{f}^{(j)}] < T, \end{cases} \quad (4)$$

где $D^{(j)}$ – двоичный сигнал на выходе, соответствующий j -му экземпляру и равный единице для класса надёжных (K_1) и нулю – для класса потенциально ненадёжных (K_2) экземпляров; $F^{(j)}$ – прогнозирующая функция, получаемая для j -го экземпляра; \mathbb{W} – оператор, определяющий вид математических операций, выполняемых над троичными сигналами $\tau_i^{(j)}$ и их весовыми коэффициентами α_i , полученными для j -го экземпляра; $i = 1, \dots, k$.

В качестве весовых коэффициентов α_i , выбираемых для j -го экземпляра в зависимости от значения его троичного сигнала τ_i , предлагается использовать условные вероятности принадлежности j -го экземпляра к классу K_1 или к классу K_2 [5, 7]. Поэтому можно говорить о весовых коэффициентах, показывающих принадлежность (близость) экземпляра к классу K_1 и принадлежность к классу K_2 при условии, что троичный сигнал τ_i этого экземпляра принял конкретное значение из числа 1, 0 или R . Указанные коэффициенты (обозначим их через $[\alpha_i^{(K_1)}(\tau_i)]$ и $[\alpha_i^{(K_2)}(\tau_i)]$ определяют по результатам обучающего эксперимента:

$$\begin{aligned} \mathfrak{b}_i^{(K_1)}(\tau_i) &= P(K_1 | \mathfrak{f} = 1) \text{ при } \mathfrak{f}^{(j)} = 1, \\ \mathfrak{b}_i^{(K_2)}(\tau_i) &= P(K_2 | \mathfrak{f} = R) \text{ при } \mathfrak{f}^{(j)} = R, \\ \mathfrak{b}_i^{(K_2)}(\tau_i) &= P(K_2 | \mathfrak{f} = 0) \text{ при } \mathfrak{f}^{(j)} = 0, \end{aligned} \quad (5)$$

где $P(K_S | \tau_i = \varepsilon)$ – условная вероятность принадлежности экземпляра к классу K_S при условии, что троичный сигнал принял значение $\tau_i = \varepsilon$ ($\varepsilon = 1; 0; R$); нижний индекс S указывает на номер класса работоспособности изделия ($S = 1; 2$); верхний индекс j при τ_i означает, что в качестве весового коэффициента должно выбираться значение, соответствующее полученному троичному сигналу для j -го экземпляра (изделия).

Вероятности $P(K_S | \tau_i = \varepsilon)$ определяют с использованием классических подходов теории вероятностей [8], обрабатывая результаты преобразования информативных параметров в троичные сигналы и принимая во внимание в обучающей выборке количество экземпляров классов K_1 и K_2 , для которых троичные сигналы приняли значения $\varepsilon = 1, 0, R$.

На основе общей модели прогнозирования (4) были предложены конкретные модели для индивидуального прогнозирования класса работоспособности экземпляра для заданной наработки. Актуальным является вопрос об эффективности моделей прогнозирования. Ответ на этот вопрос позволит сформулировать рекомендации по практическому применению моделей для индивидуального прогнозирования класса работоспособности изделий: экземпляр повышенного уровня надёжности (K_1) или потенциально ненадёжный (K_2).

Исследуемые модели прогнозирования. Для исследования на эффективность выбрано четыре модели в случае троичного преобразования информативных параметров. Математический вид прогнозирующих функций F для этих моделей приведён в таблице 1.

Таблица 1. Прогнозирующие функции моделей прогнозирования

Номер модели	Математический вид прогнозирующей функции	Суть прогнозирующей функции
1	$F^{(j)} = \prod_{i=1}^k \sigma^{(K_1)}(\Phi_i^{(j)})$	Числовой потенциал, показывающий близость (принадлежность) j -го экземпляра к классу K_1
2	$F^{(j)} = \prod_{i=1}^k \sigma^{(K_1)}(\Phi_i^{(j)}) - \prod_{i=1}^k \sigma^{(K_2)}(\Phi_i^{(j)})$	Сравнение числовых потенциалов о близости j -го экземпляра к классам K_1 и K_2
3	$F^{(j)} = \frac{\prod_{i=1}^k \sigma^{(K_1)}(\Phi_i^{(j)})}{\prod_{i=1}^k \sigma^{(K_2)}(\Phi_i^{(j)})}$	Отношение числовых потенциалов о близости j -го экземпляра к классу K_1 и к классу K_2
4	$F^{(j)} = \prod_{i=1}^k \log_2 \frac{P(K_1 \Phi_i^{(j)})}{P(K_1)} - \prod_{i=1}^k \log_2 \frac{P(K_2 \Phi_i^{(j)})}{P(K_2)}$	Количество частной информации о классе K_1 и классе K_2 , содержащейся в полученном наборе троичных сигналов τ_1, \dots, τ_k j -го экземпляра

Обозначения параметров, указанных в таблице 1:

$F^{(j)}$ – значение прогнозирующей функции, соответствующее j -му экземпляру; k – число информативных параметров, используемых для индивидуального прогнозирования класса работоспособности; $\sigma^{(K_s)}(\Phi_i^{(j)})$ – весовой коэффициент, определяемый по выражению (5), выбираемый с учётом значения троичного сигнала τ_i , полученного для j -го экземпляра ($i = 1, \dots, k$); $P(K_s|\Phi_i^{(j)})$ – условная вероятность принадлежности j -го экземпляра к классу K_s ($S = 1; 2$) при условии, что по результатам преобразования i -го информативного параметра с использованием алгоритмов (2) или (3) получен троичный сигнал $\tau_i^{(j)}$; $i = 1, \dots, k$; $P(K_s)$ – начальная вероятность принадлежности экземпляра к классу работоспособности K_s ($S = 1; 2$).

Условия статистического имитационного моделирования обучающего эксперимента. На примере биполярных транзисторов большой мощности типа КТ872А и полевых транзисторов типа КП744А были исследованы статистические закономерности информативных параметров [6, 7]. В таблице 2 в качестве иллюстрации приведены статистические характеристики информативных параметров, используемых для биполярных транзисторов большой мощности типа КТ872А. Пояснения электрических параметров, указанных в таблице 2:

- $h_{21Э}$ – статический коэффициент передачи тока в схеме с общим эмиттером;
- $I_{КЭ0}$ – обратный ток коллектора;
- $U_{КЭнас}$ – напряжение насыщения коллектор-эмиттер.

Таблица 2. Статистические характеристики информативных параметров

Характеристика	Информативный параметр					
	$x_1 \rightarrow h_{21Э}$		$x_2 \rightarrow I_{кЭ0}, \text{мкА}$		$x_3 \rightarrow U_{кЭнас}, \text{мВ}$	
	Класс K_1	Класс K_2	Класс K_1	Класс K_2	Класс K_1	Класс K_2
Минимальное значение, x_{\min}	2,97	2,68	0,15	0,16	396	392
Максимальное значение, x_{\max}	4,00	3,95	0,26	0,31	723	1395
Среднее значение, m	3,51	3,22	0,209	0,233	521,1	674,5
Стандартное отклонение, σ	0,267	0,317	0,0193	0,0293	71,15	221,4
Коэффициент вариации, ν	0,075	0,092	0,132	0,093	0,125	0,323
Отношение σ_2/σ_1	1,14		1,52		3,15	
Отношение ν_2/ν_1	1,24		1,36		2,44	

Для выполнения статистического имитационного моделирования обучающего эксперимента использовано три информативных параметра (обозначены через x_1 , x_2 и x_3). При выборе условий моделирования информативных параметров приняты во внимание экспериментально наблюдаемые расхождения между центрами классов m_1 и m_2 (средними значениями информативных параметров в классе K_1 и K_2), которые составляли до 30 %. Средние значения (математические отклонения) каждого из трёх рассматриваемых информативных параметров x_1 , x_2 и x_3 для экземпляров класса K_1 условно приняты, как $m_1 = 10$ единиц. С учётом этого, а также статистических характеристик информативных параметров, определены условия моделирования информативных параметров x_1 , x_2 и x_3 в предположении, что они имеют законы распределения, близкие к нормальным, а коэффициенты парной корреляции между информативным параметром x_i и номером класса работоспособности ИЭТ S ($S = 1, 2$), а также между самими информативными параметрами x_i соответствуют уровню «заметная корреляция» по шкале Чеддока, согласно которой модуль коэффициентов парной корреляции имеет значения $|r| = 0,5 \dots 0,7$ (таблица 3).

Таблица 3. Условия моделирования информативных параметров

Класс работоспособности ИЭТ	Характеристика информативного параметра	Значение характеристики, используемое при моделировании информативного параметра		
		x_1	x_2	x_3
K_1	m_1	10	10	10
	σ_1	0,6	0,9	1,3
K_2	m_2	11	12	13
	σ_2	0,75	1,35	2,6

Фрагмент обучающего эксперимента, смоделированного на компьютере с использованием приложения *Microsoft Excel*, показан на рисунке 2. Объём обучающей выборки выбран в количестве 5000 экземпляров, из них для заданной наработки количество

надёжных экземпляров (класс K_1) составляет $n_1 = 3000$ и потенциально ненадёжных (класс K_2) – $n_2 = 2000$.

Номер в классе	x_1	x_2	x_3	Номер класса, S
1	10,125	10,321	11,450	1
2	9,855	8,899	9,597	1
3	9,792	10,383	9,163	1
4	10,674	10,157	9,583	1
5	10,337	9,377	11,333	1
6	9,097	9,486	9,317	1
7	11,465	10,681	8,779	1
8	10,603	10,134	10,891	1
9	9,326	10,395	11,579	1
10	9,134	10,574	9,399	1

Номер в классе	x_1	x_2	x_3	Номер класса, S
1	11,209	11,517	14,502	2
2	12,274	13,969	15,661	2
3	10,715	12,265	7,027	2
4	11,869	11,406	12,639	2
5	11,213	10,775	10,179	2
6	11,219	10,613	11,243	2
7	9,907	12,372	11,596	2
8	11,052	13,134	14,068	2
9	10,998	9,980	14,176	2
10	11,210	12,626	18,057	2

2991	10,097	9,670	9,564	1
2992	9,213	9,709	10,146	1
2993	9,171	11,448	11,239	1
2994	9,590	10,162	8,730	1
2995	9,616	10,082	9,724	1
2996	10,034	12,367	10,665	1
2997	10,690	9,832	10,368	1
2998	10,437	10,226	11,939	1
2999	10,252	10,054	9,940	1
3000	10,011	9,796	10,332	1
min	4,278	6,793	5,048	
max	11,853	13,250	17,242	
m1	9,997	9,993	10,012	
$\sigma 1$	0,618	0,919	1,354	

1991	12,174	11,840	15,947	2
1992	10,335	10,727	12,456	2
1993	9,850	11,382	14,033	2
1994	10,503	11,999	9,923	2
1995	11,513	13,723	11,665	2
1996	12,448	14,685	10,818	2
1997	11,775	11,061	9,590	2
1998	9,732	11,888	11,814	2
1999	10,328	12,539	12,370	2
2000	10,959	10,959	15,690	2
min	8,850	7,877	3,012	
max	13,303	17,051	21,486	
m2	11,003	12,049	13,146	
$\sigma 2$	0,772	1,339	2,635	

Рисунок 2. Фрагмент результатов обучающего эксперимента

Используя результаты статистического имитационного моделирования обучающего эксперимента, для каждого информативного параметра были определены средние значения m_1 и m_2 для экземпляров класса K_1 и класса K_2 и по алгоритму (2) выполнено преобразование информативных параметров в троичные сигналы 1, 0 или R , а затем на основе анализа результатов преобразования найдены весовые коэффициенты, представляющие собой условные вероятности $P(K_S|\tau_i = \varepsilon)$ ($\varepsilon = 1; 0; R$); $S = 1; 2$ (таблица 4).

Таблица 4. Весовые коэффициенты троичных сигналов

Весовой коэффициент: вероятность	Значение вероятности для τ_i :			Весовой коэффициент: вероятность	Значение вероятности для τ_i :		
	τ_1	τ_2	τ_3		τ_1	τ_2	τ_3
$P(K_1 1)$	0,885	0,931	0,867	$P(K_2 1)$	0,115	0,069	0,133
$P(K_1 0)$	0,123	0,036	0,033	$P(K_2 0)$	0,877	0,964	0,967
$P(K_1 R)$	0,634	0,623	0,652	$P(K_2 R)$	0,366	0,377	0,348

С учётом математического вида прогнозирующей функции F были определены её значения для исследуемых моделей, указанных в таблице 1 под номерами 1–4. В таблице 5 в качестве примера приведены значения F для первых 20-и экземпляров класса K_1 смоделированной обучающей выборки.

В качестве критерия оценки эффективности моделей прогнозирования использовалась вероятность принятия правильных решений $P_{\text{прав}}$, определяемая по формуле [1, 2]

$$P_{\text{прав}} = \frac{n_{11} + n_{22}}{n}, \quad (6)$$

где n_{11} – количество правильно распознанных с использованием модели прогнозирования экземпляров класса K_1 в обучающей или контрольной выборке; n_{22} – количество правильно распознанных с использованием модели прогнозирования экземпляров класса K_2 в обучающей или контрольной выборке; n – объём выборки (обучающей или контрольной).

Таблица 5. Значения прогнозирующей функции для исследуемых моделей

Номер в классе	Информативный параметр			Троичный сигнал			Прогнозирующая функция F для модели (см. таблицу 1):			
	x_1	x_2	x_3	τ_1	τ_2	τ_3	1	2	3	4
1	10,125	10,321	11,450	R	R	R	1,908	0,817	1,748	0,664
2	9,855	8,899	9,597	1	1	1	2,683	2,366	8,470	7,650
3	9,792	10,383	9,163	1	R	1	2,375	1,750	3,798	4,620
4	10,674	10,157	9,583	R	R	1	2,124	1,247	2,423	2,464
5	10,337	9,377	11,333	R	1	R	2,217	1,433	2,830	3,694
6	9,097	9,486	9,317	1	1	1	2,683	2,366	8,470	7,650
7	11,465	10,681	8,779	0	R	1	1,613	0,225	1,162	-1,160
8	10,603	10,134	10,891	R	R	R	1,908	0,817	1,748	0,664
9	9,326	10,395	11,579	1	R	R	2,160	1,319	2,570	2,820
10	9,134	10,574	9,399	1	R	1	2,375	1,750	3,798	4,620
11	9,771	9,943	13,369	1	1	0	1,849	0,697	1,606	0,050
12	9,202	10,040	10,977	1	R	R	2,160	1,319	2,570	2,820
13	11,126	11,112	8,962	0	R	1	1,613	0,225	1,162	-1,160
14	9,243	11,800	11,207	1	R	R	2,160	1,319	2,570	2,820
15	10,327	9,887	11,312	R	1	R	2,217	1,433	2,830	3,694
16	9,935	9,503	11,290	1	1	R	2,468	1,936	4,639	5,850
17	9,683	9,078	11,160	1	1	R	2,468	1,936	4,639	5,850
18	10,281	9,854	8,295	R	1	1	2,432	1,864	4,281	5,494
19	10,452	9,808	10,768	R	1	R	2,217	1,433	2,830	3,694
20	9,859	9,173	9,440	1	1	1	2,683	2,366	8,470	7,650

Результаты исследований. В таблице 6 приведены значения вероятностей $P_{\text{прав}}$, полученные путём применения построенных моделей (с использованием данных смоделированного обучающего эксперимента) для индивидуального прогнозирования класса работоспособности (K_1 или K_2) экземпляров этой же обучающей выборки (5000 экземпляров). Указываются также усреднённые значения $P_{\text{прав}}$, полученные по результатам применения построенных по обучающей выборке исследуемых моделей к пяти смоделированным контрольным выборкам, каждая из которых имела объём 5000 экземпляров. Кроме того, приводятся значения $P_{\text{прав}}$, полученные по моделям прогнозирования, построенным с учётом преобразования информативных параметров в двоичные сигналы (1 и 0). Эти данные приведены для модели базового метода пороговой логики [1] и модели, обеспечивающей максимальное значение вероятности $P_{\text{прав}}$ из числа исследуемых моделей прогнозирования, использующих двоичное преобразование информативных параметров.

Таблица 6. Эффективность моделей прогнозирования по результатам статистического имитационного моделирования

Номер $F^{(j)}$ в таблице 1	Вид кодового сигнала	Описание прогнозирующей функции $F^{(j)}$	Значение порога разделения классов	Вероятность правильных прогнозов класса работоспособности	
				обучающая выборка	контрольная выборка
–	Двоичный	Базовый метод пороговой логики, использование близости к классу K_1 , коды: 1, 0	1,433*	0,894	0,871
–	Двоичный	Использование частной информации о классах K_1 и K_2	0 бит	0,906	0,894
1	Троичный	Аналог базового метода пороговой логики, использование близости к классу K_1 , коды: 1, 0, R	1,428*	0,906	0,894
2	Троичный	Разность показателей близости к классам K_1 и K_2	0	0,894	0,890
3	Троичный	Отношение показателей близости к классам K_1 и K_2	0	0,901	0,894
4	Троичный	Использование частной информации о классах K_1 и K_2	0 бит	0,913	0,906

* – Пороги найдены из условия $P_{\text{прав}} \rightarrow \max$.

Заключение. Как следует из результатов, приведённых в таблице 6, модели прогнозирования, использующие суммарное значение частной информации о классе K_1 и классе K_2 , получаемой от набора двоичных (1 или 0) и троичных (1, 0 или R), в рамках своего вида преобразования информативных параметров обеспечивают большее значение вероятности принятия правильных решений, т.е. меньшие ошибки прогнозирования класса работоспособности изделий, нежели другие модели. С использованием результатов индивидуального прогнозирования, полученным при рассмотрении обучающей и контрольных выборок установлено, что модель прогнозирования, использующая суммарное значение частной информации о классе K_1 и классе K_2 , в случае троичного преобразования информативных параметров обеспечивает ошибки прогнозирования на 10...20 % меньшие, нежели аналогичная модель в случае двоичного преобразования информативных параметров. Поэтому эта модель рекомендуется для использования при решении практических задач индивидуального прогнозирования класса работоспособности изделий электронной техники.

Список литературы

- [1] Боровиков, С. М. Статистическое прогнозирование для отбраковки потенциально ненадёжных изделий электронной техники: монография / С. М. Боровиков. – М. : Новое знание, 2013. – 343 с.
- [2] Пестряков, В. Б. Индивидуальное прогнозирование состояния РЭА с использованием теории распознавания : учеб. пособие / В. Б. Пестряков, В. В. Андреева. – Куйбышев : Изд-во КуАИ, 1980. – 88 с.
- [3] Дертоузос, М. Пороговая логика : пер. с англ. / М. Дертоузос. – М. : Мир, 1967. – 343 с.
- [4] Метод прогнозирования надёжности изделий электронной техники / С. М. Боровиков [и др.] // Доклады НАН Беларуси. – 2006. – Т. 50, № 4. – С. 105-109.
- [5] Прогнозирование надёжности изделий электронной техники методом пороговой логики / С. М. Боровиков [и др.] // Доклады БГУИР : электроника, материалы, технологии, информатика. – 2006. – № 2(14). – С. 49-56.

[6]. Прогнозирование класса надёжности изделий электронной техники методом преобразования информативных параметров в дискретный код / В. О. Казючиц [и др.] // Доклады ТУСУР. – 2023. – Т. 26, № 1. – С. 91–97. DOI: 10.21293/1818-0442-2023-26-1-91-97.

[7] Казючиц В. О., Боровиков С. М., Шнейдеров Е. Н. // Доклады БГУИР. – 2022. – Т. 20, № 1. – С. 92- 100. – DOI : <http://dx.doi.org/10.35596/1729-7648-2022-20-1-92-100>.

[8] Вентцель Е. С. Теория вероятностей : учеб. для вузов. – 10-е изд. стер. – М. : Высш. шк., 2006. – 575.

STATISTICAL SIMULATION MODELING IN THE RESEARCH OF THE EFFECTIVENESS OF MODELS FOR PREDICTING PRODUCT RELIABILITY BY INFORMATIVE PARAMETERS

S.M. Borovikov

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

Abstract. The application of statistical simulation modeling is considered to study the effectiveness of models for predicting the reliability class of electronic equipment products for a given operating time based on the results of monitoring the informative parameters of products at the initial point in time. A feature of the forecasting models under study is the transformation of informative parameters into discrete binary or ternary code. Several ways to build models based on the use of binary or ternary signals are considered. The models are obtained by processing data from a computer-simulated training experiment, and the effectiveness of the models is determined by applying them both to instances of the training set and to instances of the simulated control samples.

Based on the research, recommendations are given on the use of models to solve practical problems of predicting the reliability class of products for given operating times.

Keywords: informative parameters, conversion to code signals, product reliability class, forecasting models.

УДК 502.175

СИСТЕМА АВТОМАТИЧЕСКОГО ОПОВЕЩЕНИЯ И КОНТРОЛЯ КОНЦЕНТРАЦИИ МЕТАНА И УГЛЕКИСЛОГО ГАЗА В ВОЗДУХЕ ПРОИЗВОДСТВЕННОГО ПОМЕЩЕНИЯ



И.А. Телеш

Доцент кафедры инженерной психологии и эргономики БГУИР, кандидат географических наук, доцент
tia32@bsuir.by



Е.А. Азаров

Студент кафедры информатики факультета компьютерных систем и систем БГУИР-
azarov@gmail.com



Д.С. Шумский

Студент кафедры информатики факультета компьютерных систем и систем БГУИР
dan.shumski@gmail.com



А.В. Горгун

Студент кафедры информатики факультета компьютерных систем и систем БГУИР
gorgun@gmail.com

И.А. Телеш

Окончила Белорусский государственный педагогический университет имени М. Танка. Область научных интересов связана с применением информационных технологий в техносферной безопасности, организацией учебного и научно-исследовательского процессов в техническом университете.

Е.А. Азаров

Студент 2 курса специальности «Информатика и технологии программирования» кафедры информатики факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники.

Д.С. Шумский

Студент 2 курса специальности «Информатика и технологии программирования» кафедры информатики факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники.

А.В. Горгун

Студент 2 курса специальности «Информатика и технологии программирования» кафедры информатики факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники.

Аннотация. Разработана система автоматического оповещения, отображающая концентрацию метана и углекислого газа в воздухе производственного помещения. При предельных показаниях анализируемых газов экстренные уведомления отправляются в мессенджер. Описана архитектура IoT, модули и датчики, используемые для разработки устройства.

Ключевые слова: система автоматического извещения и контроля, IoT, C# ASP.NET, Entity Framework, SQLite, ESP-32, JavaScript.

Введение. За состоянием производственных факторов на рабочих местах в соответствии с санитарными нормами и правилами организуется контроль. Контроль за количественным содержанием вредных веществ в воздухе рабочей зоны должен

проводиться по величинам максимально разовых и (или) среднесменных значений загрязнения воздушной среды при сравнении их с ПДК_{мр} и ПДК_{сс}, установленными гигиеническими нормативами [1].

В условиях производства и местах пребывания рабочих проводится контроль за количественным содержанием вредных веществ в воздухе рабочей зоны, среди которых углекислый газ и метан. Углекислый газ в высоких концентрациях является токсичным, его относят к удушающим газам и IV классу опасности [1]. При повышении концентрации углекислого газа в воздухе (0,15 % - 0,2 %) у человека возникает ощущение духоты: общий дискомфорт, слабость, сонливость, головная боль, снижение концентрации внимания, и как следствие снижается работоспособность.

Помимо углекислого газа на производстве встречается и газ метан, который проникает в организм через органы дыхания, заменяет кислород и разносится кровью. При высокой концентрации метана в закрытых помещениях снижается количество вдыхаемого кислорода, подавляется дыхание, поражается центральная нервная система, наступает сильное кислородное голодание и возможно удушающее воздействие.

В связи с этим в производственных условиях и в местах пребывания рабочих возникает необходимость в обнаружении превышенной концентрации газообразных веществ, а также газов.

Целью работы является разработка устройства автоматического оповещения и контроля уровня метана и углекислого газа в воздухе производственного помещения, в котором возможно возникновение чрезвычайной ситуации.

Основная часть. В ходе работы произведена сборка устройства и разработка системы автоматического оповещения и контроля по измерению текущего уровня метана и углекислого газа с последующим отображением полученной информации.

Архитектура системы автоматического оповещения и контроля углекислого газа и метана состоит из следующих блоков:

- устройство сбора и отправки на сервер данных датчиков о текущем уровне метана и углекислого газа;
- серверное приложение.

Устройство сбора показаний датчиков основано на микроконтроллере *ESP-32*, при помощи которого осуществляется передача показаний на удаленный сервер каждые две минуты. Выбор микроконтроллера *ESP-32* для разработанного устройства позволил реализовать непрерывный сбор данных, а в случае опасного для здоровья человека, немедленную отправку тревожного сигнала о превышении предельной концентрации уровня метана и углекислого газа в воздухе производственного помещения.

Для сбора показаний на устройстве расположено два датчика, определяющие концентрацию определенного газа в воздухе: *MQ-135* (углекислый газ CO_2) и *MQ-4* (метан CH_4), а также датчики *Dallas 18B20* (температура воздуха в помещении) и *DHT-11* (влажность воздуха в помещении). При помощи датчиков температуры и влажности происходит корректировка показаний концентрации газов.

На рисунке 1 изображена схема подключения радиокомпонентов в разработанном устройстве для автоматического оповещения, о превышении концентрации углекислого газа и метана в воздухе производственного помещения.

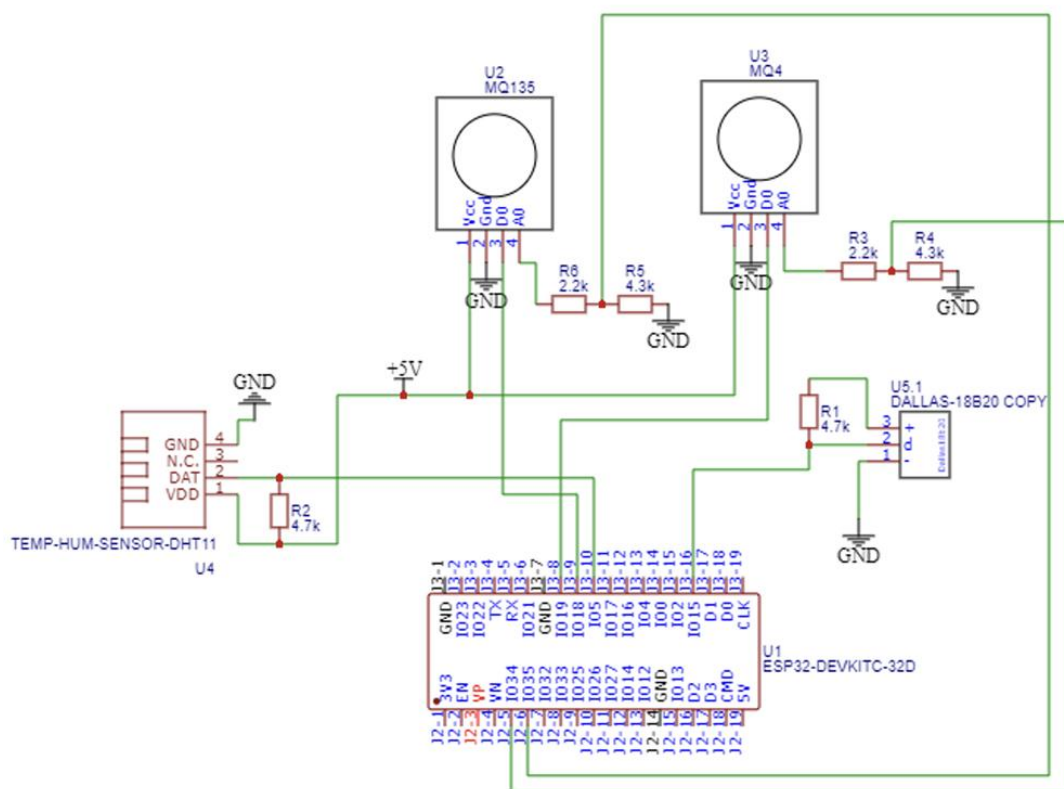


Рисунок 1. Электрическая схема разработанного устройства

На рисунке 2 показана сборка разработанного устройства на макетной плате.

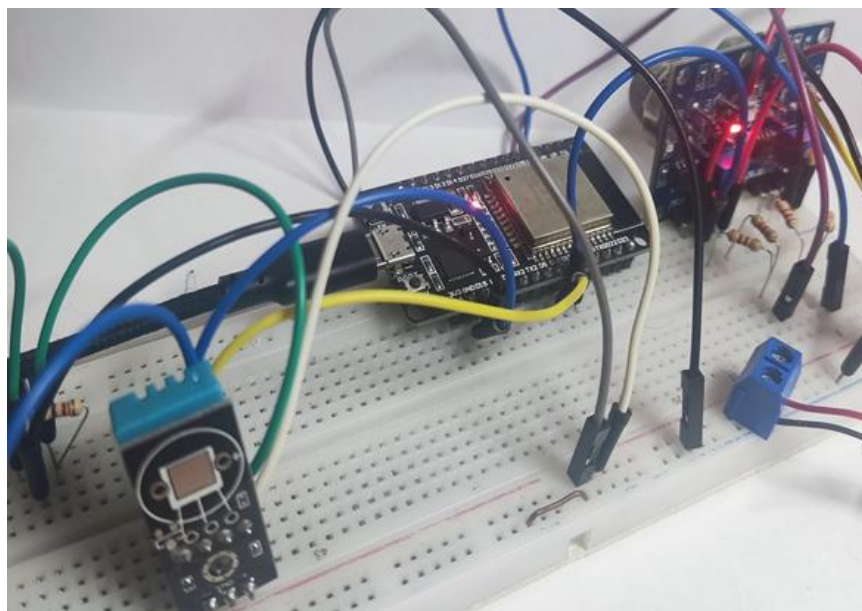


Рисунок 2. Макетная плата и схема устройства на ней

После тестирования работоспособности разработанного устройства стало возможным компактное размещение радиодеталей на монтажной плате, как показано на рисунке 3.

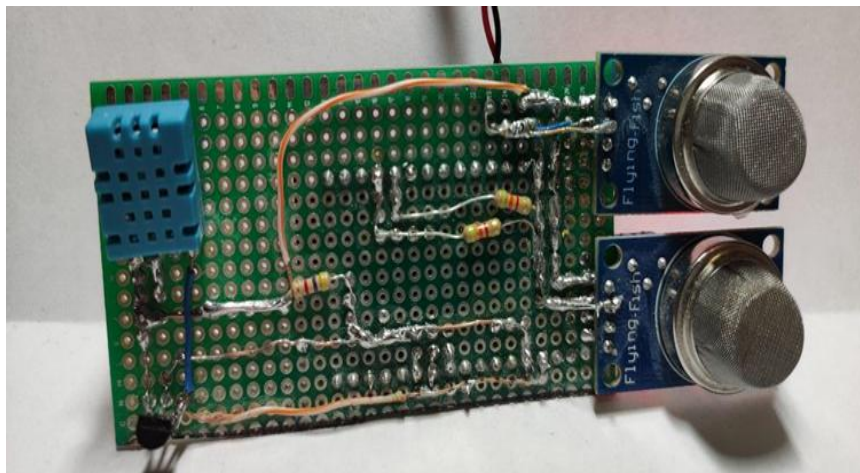


Рисунок 3. Размещение радиодеталей устройства на монтажной плате

Серверное приложение разработано с использованием технологий *C# ASP.NET*, *Entity Framework*, *SQLite*. Серверное приложение принимает и обрабатывает показания датчиков, определяющих концентрацию углекислого газа и метана в воздухе, сохраняет их в базу данных для последующего отображения в виде графиков. В случае получения данных, превышающих предельно допустимые значения, технологии *Docker* серверное приложение развернуто на хостинге. Интерфейс программирования приложения (*API*) для работы с сервером представлен в таблице 1.

Таблица 1. Структура *API* для работы с сервером

Эндпоинт	Описание	Тело запроса	Тело ответа
<i>POST /api/rawdata/pushextra</i>	Экстренное оповещение	<i>JSON: {"statusCode":0}</i>	Код статуса: 200
<i>POST /api/rawdata/push</i>	Добавление данных на сервер	<i>JSON: [{"date": 0, "temperatureC": 0, "humidity": 0, "cO2": 0, "cH4": 0}]</i>	Код статуса: 200
<i>GET /api/rawdata/getforday/{день}-{номер месяца}-{год}</i>	Получение данных за определенный день		<i>JSON: [{"date": 0, "temperatureC": 0, "humidity": 0, "cO2": 0, "cH4": 0}]</i>
<i>GET /home</i>	Получение html страницы с графиками		

На рисунках 4 и 5 в виде графиков отображены данные с датчиков устройства о концентрациях углекислого газа и метана в воздухе производственного помещения в случае возникновения чрезвычайной ситуации, при этом учитывая температуру и влажность воздуха в течение суток.

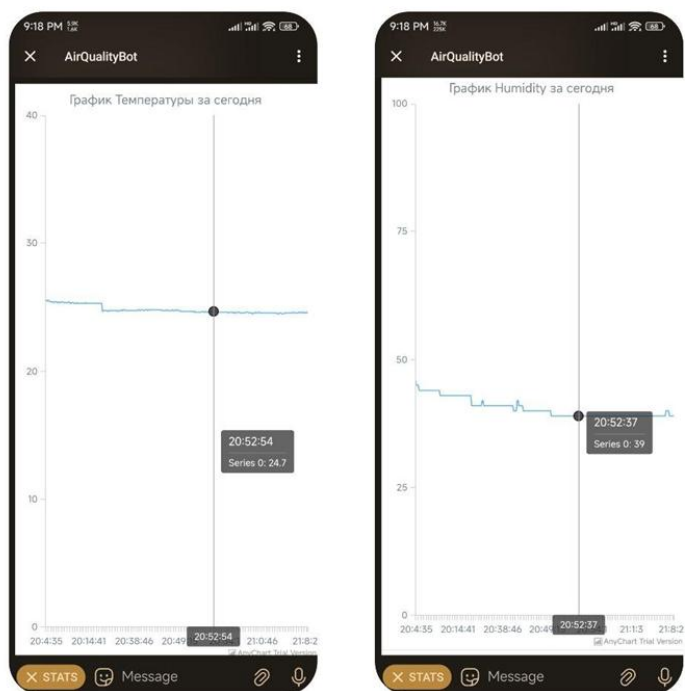


Рисунок 4. Отображение сведений с датчиков устройства о температуре и влажности в производственном помещении

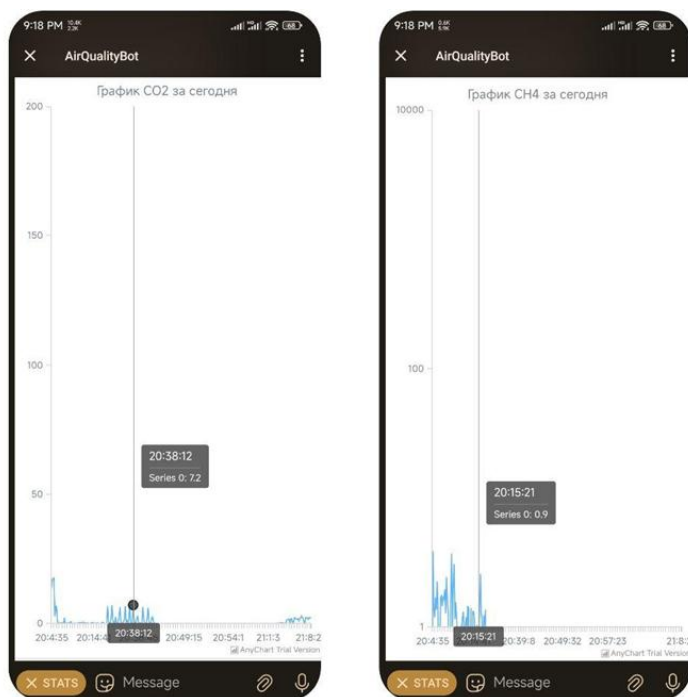


Рисунок 5. Сообщение о превышении концентрации углекислого газа и метана в случае чрезвычайной ситуации

На рисунке 6 показано экстренное оповещение о чрезвычайной ситуации в случае превышения показаний датчиков.

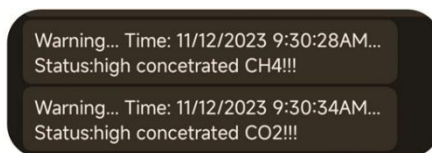


Рисунок 6. Оповещение о превышении показаний одного из датчиков на разработанном устройстве

Заключение. Разработана система автоматического оповещения и контроля уровня концентрации метана и углекислого газа в воздухе производственного помещения, в котором возможно возникновение чрезвычайной ситуации. Для проектирования устройства применены информационные технологии: *ASP.NET*, *ESP-32* и *JavaScript*. Система автоматического оповещения и контроля уровня концентрации метана и углекислого газа в воздухе предоставляет данные в виде наглядных графиков, используя которые возможно оценить степень загрязнения воздуха в производственном помещении при определенном местоположении устройства.

Список литературы

- [1]. ГН-1 от 11.10.2017 № 92 Гигиенический норматив "Предельно допустимые концентрации вредных веществ в воздухе рабочей зоны".
- [2]. Arduino Documentation [Электронный ресурс].- 2023.- Режим доступа : <https://docs.arduino.cc/>.
- [3]. MQ 135 Datasheet [Электронный ресурс]. - 2023. - Режим доступа:https://www.electronicoscaldas.com/datasheet/MQ-135_Hanwei.pdf.
- [4]. ESP 32 Documentattion [Электронный ресурс] - 2023.- Режим доступа :<https://espressif-docs.readthedocs-hosted.com/>.
- [5]. DHT 11 Datasheet [Электронный ресурс].- 2023.Режим доступа: https://www.electronicoscaldas.com/datasheet/DHT11_Aosong.pdf.
- [6]. ASP.NET Documentation [Электронный ресурс].- 2023.- Режим доступа : <https://learn.microsoft.com/en-us/aspnet/core/>.

Авторский вклад

Авторы внесли равноценный вклад в написании статьи

METHANE AND CARBON DIOXIDE CONCENTRATION AUTOMATIC NOTIFICATION AND CONTROL SYSTEM IN THE AIR OF A PRODUCTION PREMISES

I.A. Telesh
Associate Professor of
the Department of
Engineering Psychology
and Ergonomics of
BSUIR, Candidate of
Geographical Sciences,
Associate Professor

E.A. Azarov
Student of the
Department of
Informatics, Faculty of
Computer Systems and
Systems of BSUIR

D.S. Shumsky
Student of the
Department of
Informatics, Faculty of
Computer Systems and
Systems of BSUIR

A.V. Gorgun
Student of the
Department of
Informatics, Faculty of
Computer Systems and
Systems of BSUIR

Abstract. An automatic notification and control system has been developed that displays the concentration of methane and carbon dioxide in the air of the production premises. If the readings of the analyzed gases are extreme, emergency notifications are sent to the messenger. The IoT architecture, modules and sensors used to develop the device are described.

Keywords: automatic notification and control system, IoT, C# ASP.NET, Entity Framework, SQLite, ESP-32, JavaScript.

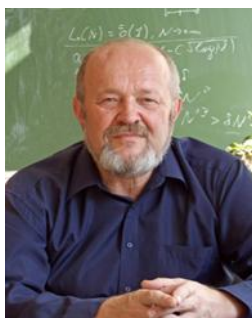
УДК 004.8

АКТУАЛЬНЫЕ ВОПРОСЫ ПРЕПОДАВАНИЯ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, НАУКИ О ДАННЫХ И МАШИННОГО ОБУЧЕНИЯ



М.В. Спринджук

Доцент кафедры экономической информатики БГУИР, кандидат технических наук, лектор, Старший научный сотрудник ОИПИ НАН Беларуси
stepanenkomatvei@yandex.ru



В.И. Берник,

доктор физ.-мат. наук, профессор, главный научный сотрудник Института математики НАН Беларуси



В.С. Филинук

магистрант кафедры ИИТ БГУИР

М.В. Спринджук

Окончил аспирантуру БГУИР. Область научных интересов связана с разработкой алгоритмов, моделей и программного обеспечения с использованием технологий искусственного интеллекта.

В.И. Берник

Окончил БГУ, специалист в области теории чисел и прикладной математики.

В.С. Филинук

Окончила Белорусский государственный экономический университет. Область интересов связана с исследованием применения искусственного интеллекта в целях поддержки принятия решения в банковской сфере. Обучается в магистратуре кафедры Искусственного интеллекта БГУИРа.

Аннотация. Приводятся идеи и концепции для развития организации и практики преподавания науки о данных и сопряженных технологий искусственного интеллекта.

Ключевые слова: высшее техническое образование, искусственный интеллект, наука о данных, большие данные, машинное обучение, педагогическая деятельность, качество образования.

Введение. Цель современного высшего технического образования – подготовить конкурентноспособных специалистов за сравнительно короткий временной период обучения бакалавра и магистра. **Актуальность темы** организации преподавания технологий искусственного интеллект возросла по причине снижения цены персональных и мобильных компьютеров, интернета, мультимедийных устройств и по мере развития математического и алгоритмического аппарата и доступных программных библиотек.



Рисунок 1. Сферы применения и междисциплинарности искусственного интеллекта [4].
ООП- объектно-ориентированное программирование

Материалы и методы. Для поиска релевантной научной литературы были использованы ресурсы программного обеспечения *EndNote 21* [<https://endnote.com/>] и вебсайт *Google scholar* [<https://scholar.google.com/>].

Результаты и обсуждение. Высшее образование как совокупность усвоенной информации, которая передается за определенное количество часов, требует соответствия следующих критериев:

- 1 Современности и актуальности.
- 2 Оптимальности по объему, структуре и составу.
- 3 Сбалансированной по степени сложности и доступности к усвоению.
- 4 Универсальностью, соответствием с информацией крупнейших зарубежных ВУЗов (высших учебных заведений).
- 5 Научемкостью и последовательностью в преподавании.

Наука о данных (*data science*) и сопряженные с ней технологии искусственного интеллекта и машинного обучения, в частности, требуют от учащегося интеграции знаний, накопленных еще со школьных курсов математики и информатики и усвоения множества предметов в стенах университета.

Обзор релевантной научно-педагогической литературы показывает, что для преподавания науки о данных, технологий *Больших данных* и *Искусственного интеллекта* (рисунок 1) [1,2,3,5,6,7], рекомендовано использовать следующие подходы, технические средства и методики:

- программное обеспечение для удаленных лекций и конференций;
- облачные технологии для хранения и передачи больших объемов информации;
- планшеты и другие мобильные устройства для просмотра кодов, результатов вычислений, для согласования организационных вопросов и т.п.;
- проведение конкурсов-олимпиад по информатике, программированию, математике для профориентации учащихся и для поощрения талантливых обучающихся;
- поощрение научных публикаций студентов для повышения престижа науки и социально-экономического статуса молодого ученого;
- открытие специализированных конкурсов для финансирования научных проектов с целью развития научных основ теории и практики применения искусственного интеллекта и науки о данных;
- привлечение научных партнеров и экспертов из дружественных стран для сотрудничества, обмена информацией, взаимного внедрения;

- приглашение лекторов ученых и практиков-специалистов из разных университетов и государств;
- проведение семинаров и конференций компаниями и предприятиями-работодателями для подбора сотрудников;
- обучение и поощрение самостоятельной работы учащихся и слушателей курсов, как основного подхода к усовершенствованию индивидуального уровня образования;
- разработка мультимедийных образовательных ресурсов для самообучения навыкам работы с вычислительной техникой и программным обеспечением.

Заключение. Таким образом, преподавание науки о данных, технологиям анализа и визуализации больших данных и сопряженным технологиям искусственного интеллекта необходимо развивать по направлениям расширения международного сотрудничества, усовершенствования технического оснащения, оптимизации организации педагогического труда, поощрения научной активности и внедрения результатов.

Список литературы

- [1] Evangelista, I., Blesio, G., Benatti, E. Why are we not teaching machine learning at high school? A proposal. IEEE, 2018. — P. 1-6.
- [2] Gresse von Wangenheim, C., Hauck, J.C., Pacheco, F.S., Bertoneceli Bueno, M.F. Visual tools for teaching machine learning in K-12: A ten-year systematic mapping // Education and Information Technologies.— 2021.— V. 26, № 5.— P. 5733-5778.
- [3] Haack, M., Jambor, T.N. Remote laboratory for activating lecture hall teaching and distance learning. 2021. — P. 536-542.
- [4] Harris, H.D., Kiefer, S.M. The Pedagogy of Artificial Intelligence: A Survey of Faculty Who Teach Introductory AI. Citeseer, 2004. — P. 74-79.
- [5] Marques, L.S., Gresse von Wangenheim, C., Hauck, J.C. Teaching machine learning in school: A systematic mapping of the state of the art // Informatics in Education.— 2020.— V. 19, № 2.— P. 283-321.
- [6] Martins, R.M., Von Wangenheim, C.G. Findings on teaching machine learning in high school: A ten-year systematic literature review // Informatics in Education.— 2023.— V. 22, № 3.— P. 421.
- [7] Shamir, G., Levin, I. Teaching machine learning in elementary school // International Journal of Child-Computer Interaction.— 2022.— V. 31.— P. 100415. Epub.

Авторский вклад

Равный вклад авторов.

CURRENT ISSUES IN TEACHING TECHNOLOGIES OF ARTIFICIAL INTELLIGENCE, DATA SCIENCE AND MACHINE LEARNING

Sprindzuk M.V.

*PhD of Technical Sciences
Senior research scientist
Computer science lecturer
Working on position of
associate professor
Senior researcher at the UIIP,
NASB, Minsk*

Bernik V.I.

*Doctor of Physical and
Mathematical Sciences, Professor,
Institute of Mathematics of the
National Academy of Sciences of
Belarus, Minsk.*

Filinuk V.S.

*Masters degree student at the
cathedra of Intellectual
technologies, BSUIR, Minsk*

Abstract. Ideas and concepts for developing the organization and practice of teaching data science and related artificial intelligence technologies are presented.

Keywords: higher technical education, artificial intelligence, data science, big data, machine learning, pedagogical activity, quality of education.

УДК 004.891.2

ИСПОЛЬЗОВАНИЕ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АВТОМАТИЗАЦИИ БИЗНЕС-ПРОЦЕССОВ



О.Н. Шкор

Старший преподаватель кафедры экономики БГУИР
shkor@bsuir.by

О.Н. Шкор

Родилась в Минске. Закончила БПИ в 1984 году. В 2001 г. защитила магистерскую диссертацию на тему: «Использование блочно-модульной системы обучения в профессиональной ориентации школьников» по специальности 08.00.05 «Экономика и управление народным хозяйством». В 2010 г. защитила докторскую диссертацию (PhD) на тему: «Экономическое обоснование формирования транспортной логистической системы Республики Беларусь» в Международной кадровой академии (Киев). С 2014 г. по настоящее время - заместитель заведующего кафедрой по научно-исследовательской работе студентов.

Аннотация. В статье рассматриваются перспективы развития генеративного искусственного интеллекта, приводится мнение аналитиков по поводу основных трендов использования искусственного интеллекта в различных сферах. Приводится несколько потенциальных сфер применения мобильных приложений с встроенным ИИ. Анализируются возможности использования генеративного искусственного интеллекта в программировании, маркетинге, игровой индустрии.

Ключевые слова: генеративный искусственный интеллект, маркетинг, игровая индустрия, чат-боты, подкасты.

Введение. В настоящее время по всему миру наблюдается резкий рост интереса к технологиям генеративного ИИ (*GenAI*). Эти современные инновации способны создавать тексты, изображения и разнообразный контент на основе данных, используемых для обучения моделей. Ожидается, что к 2027 году объем рынка искусственного интеллекта(ИИ) достигнет 407 миллиардов долларов. С 2023 по 2030 год темпы роста ИИ составят 37,3% в год. В 2030 году выручка глобального рынка ИИ достигнет 1,8 триллиона долларов. Более 1800 компаний по всему миру разрабатывают модели естественного языка для различных сфер. К 2030 году мировой ВВП вырастет на 15,7 триллиона долларов благодаря ИИ. В 2024 году 314 миллионов человек будут использовать инструменты ИИ. К 2026 году 90% интернет-контента будет создаваться при помощи искусственного интеллекта.

IDC прогнозирует, что стремительное смещение ИТ-затрат в сторону искусственного интеллекта затронет почти все отрасли и области применения. К 2025 году 2000 крупнейших компаний мира (*G2000*) будут направлять более 40% своих основных ИТ-затрат на инициативы, связанные с ИИ. Эффективность моделей ИИ в компаниях, внедряющих прозрачность, доверие и безопасность, вырастет на 50% к 2026 году. 92% разработчиков в США уже используют инструменты ИИ для задач программирования как на работе, так и вне ее. 8,4 млрд достигнет число устройств с голосовым помощником в мире к 2024 году – это больше, чем население Земли. Предполагается что 70% новых веб- и мобильных приложений благодаря ИИ будут

разрабатываться иначе уже к 2026 году. Каждый десятый автомобиль будет беспилотным благодаря технологиям ИИ к 2030 году. \$1 млрд ежегодно экономит стриминговый сервис *Netflix*, используя машинное обучение в целях создания индивидуальных рекомендаций для пользователей. На 225% ускорила время выбора продуктов торговая площадка *Amazon* благодаря машинному обучению[1].

Реальность и прогнозы выглядят очень оптимистично. Однако, многие сотрудники компаний не скрывают своего волнения по поводу дальнейшей карьеры. И у них есть повод для беспокойства. Есть прогнозы, что от 75 млн до 375 млн человек придется сменить профессию уже к 2030 году. До 40% к 2035 году повысится производительность труда в 16 отраслях, включая обрабатывающую промышленность, благодаря внедрению технологий ИИ. 30% компаний внедряют технологии искусственного интеллекта из-за нехватки человеческих ресурсов и для автоматизации рутинных задач. А основатель *Microsoft* Билл Гейтс считает, что только 3 дня в неделю будут работать люди, когда ИИ станет достаточно развитым.

И, конечно, уже сейчас мы наблюдаем как активно внедряется ИИ не только в бизнес, но и в повседневную жизнь. Ответы на сообщения, ответы на финансовые вопросы, планирование маршрутов путешествий и публикация контента в социальных сетях стали наиболее популярными вариантами применения ИИ в 2023 году. Даже прогноз погоды нам генерирует нейронная сеть. 1 млн пользователей насчитывался у чат-бота *ChatGPT* через пять дней после запуска. И, конечно, быстрее всего нейронные сети освоили студенты, которым постоянно не хватает то времени, то знаний для написания курсовых и лабораторных. Да и преподавателям для написания уникальных лекций, статей, докладов, презентаций они тоже очень пригодились. И, если мы не освоим эти новые технологии сейчас, мы навсегда останемся на уровне ручного труда, в то время как все те, кто овладел навыками в использовании ИИ, резко повысят свою производительность. 37% компаний инвестируют в обучение и внедряют другие стимулы для более быстрого осваивания сотрудниками навыков работы с ИИ. Предполагается, что до 40% к 2035 году повысится производительность труда в 16 отраслях, включая обрабатывающую промышленность, благодаря внедрению технологий ИИ.

ИТ-сектор почувствует воздействие ИИ сильнее, чем любая другая область, поскольку почти каждая компания стремится предложить продукты и услуги, основанные на нейронных сетях, машинном обучении и больших языковых моделях. Кроме того, компании активно помогают своим клиентам во внедрении ИИ.

Генеративный ИИ приведет к изменениям во многих сферах. *IDC* предсказывает, что к 2025 году 40% сервисов, таких как оценка рисков и ИТ-операции, будут использовать инструменты *GenAI*. Эти средства позволят создавать виртуальных помощников, генерирующих человекоподобные ответы, разрабатывающих динамичные видеоигры и генерирующих синтетические данные для обучения других моделей ИИ. Аналитики рекомендуют организациям планировать, тестировать и внедрять полностью сближенные ИИ-решения, чтобы развивать новые услуги с учетом потребностей клиентов и при этом экономить на затратах. [2]

Благодаря алгоритмам машинного обучения, ИИ может анализировать огромные объемы информации и выявлять скрытые закономерности, что помогает предпринимателям принимать более обоснованные решения. Кроме того, ИИ способен улучшить опыт клиентов, предлагая персонализированные рекомендации и быстрое обслуживание. Это помогает компаниям укрепить свою репутацию и удержать лояльных клиентов.

ИИ открывает огромные возможности для анализа данных и прогнозирования поведения потребителей. С его помощью можно создавать персонализированные рекламные кампании, оптимизировать контент и улучшать взаимодействие с клиентами.

Анализ данных позволяет понять предпочтения и потребности целевой аудитории. Используя данные о поведении клиентов, можно создавать более точные и эффективные маркетинговые стратегии. Например, предсказывать, какие товары интересны определенному клиенту или, когда лучше отправлять персонализированные предложения. После внедрения ИИ чат-бота, сотрудники освобождаются от этой рутины и могут сконцентрироваться на более важных задачах. 1,5 млрд человек в мире уже используют чат-ботов. Чат-боты способствуют автоматизации продаж. Они могут проводить клиентов через весь процесс покупки, от подбора товара до оформления заказа. Это сокращает время на принятие решения и увеличивает конверсию. Использование ИИ чат-ботов в бизнесе открывает новые возможности для повышения доходов и улучшения конкурентоспособности компании. Развитие технологий в этой области будет продолжаться, и компании, которые активно внедряют инновации, смогут выйти вперед на рынке. Искусственный интеллект может помочь улучшить эти процессы продаж. Согласно исследованию, проведенному *Harvard Business Review*, компании, использующие ИИ в продажах, смогли увеличить количество лидов более чем на 50%, сократить время звонков на 60-70% и добиться снижения затрат на 40-60% [3]. Для бизнесов это огромное преимущество: чат-боты работают круглосуточно, не тратят время на отдых или больничные, их можно легко масштабировать и они обеспечивают однородное обслуживание для всех клиентов. По прогнозам к концу 2024 года более половины всех бизнесов будут использовать чат-боты с искусственным интеллектом в своей работе.

И еще одно направление в бизнесе, которое также сейчас в разработке ИИ. Это разработка мобильных приложений, игр и подкаст.

По прогнозам Института инженеров электротехники и электроники (*IEEE*), уже в 2024 году приложения на основе искусственного интеллекта будут использоваться в самых разных областях. ИИ будет помогать оптимизировать данные, решать сложные задачи и принимать решения с человеческой точностью. *IEEE* называет:

- 1 выявление уязвимостей кибербезопасности в реальном времени и предотвращение атак;
- 2 улучшение эффективности автоматизации цепочек поставок и работы складов;
- 3 помощь и ускорение разработки программного обеспечения;
- 4 автоматизация обслуживания клиентов;
- 5 ускорение проверки соискателей, вакансий, подбор и найм персонала;
- 6 ускорение исследований заболеваний и разработка лекарств;
- 7 автоматизация и стабилизация источников электроснабжения.

Аналитики прогнозируют, что в 2024 году 314 миллионов человек будут использовать инструменты ИИ, а в 2026 году 90% интернет-контента будет создаваться при помощи искусственного интеллекта. Уже сейчас мы каждый день видим новые тренинги, вебинары и обучающие курсы, где нам предлагают за неделю создать с помощью *GenAI* свое мобильное приложение, научиться программировать на *Python* в связке с *ChatGPT*, создавать и вести не просто подкасты, а полноценное нейросетевое шоу, где контент создается исключительно с помощью ИИ. И это не обычная озвучка блокпостов, а уникальное сочетание автоматизированного анализа, креативной обработки и генерации контента.

Нельзя упустить из виду и игровую индустрию. В ней разработчики игр с искусственным интеллектом являются важным активом в игровой индустрии, поскольку они обладают специальными навыками для создания и тонкой настройки алгоритмов, которые оживляют динамичные игровые процессы. Благодаря включению технологий искусственного интеллекта в игровую механику происходит захватывающая трансформация, поскольку игры становятся способными приспосабливаться к выбору и

склонностям отдельных игроков. Эта динамическая интеграция приводит к персонализированным и постоянно развивающимся игровым приключениям. Разработчики игр с искусственным интеллектом постоянно исследуют новые горизонты игровых инноваций, увлекая игроков и обогащая игровую вселенную в целом. Это происходит от создания интеллектуальных неигровых персонажей (NPC) и использования процедурной генерации контента до повышения аутентичности виртуальных миров. Аналитики данных ИИ позволяют разработчикам игр и маркетологам настраивать свои предложения и рекламные акции для определенных сегментов аудитории, что в конечном итоге приводит к более увлекательным игровым процессам и более высокому уровню удовлетворенности и удержания игроков за счет расшифровки предпочтений игроков, моделей взаимодействия и поведения в игре.

Конечно, многие обеспокоены таким стремительным вхождением в нашу жизнь ИИ. Например, более 75% потребителей контента обеспокоены опасностью распространения дезинформации с помощью ИИ. 56% компаний называют неточность работы моделей главным риском при внедрении ИИ. 77% респондентов опасаются, что ИИ может привести к потере рабочих мест в ближайшем будущем. При этом 300 млн работников по всему миру может вытеснить искусственный интеллект по мере своего развития. От 75 млн до 375 млн человек придется сменить профессию уже к 2030 году. В 14 раз чаще будут увольнять работников низкооплачиваемых профессий в США к 2030 году из-за автоматизации их обязанностей. Но при этом 97 млн рабочих мест создаст ИИ к 2025 году. И это, безусловно, необходимо понимать уже сейчас и будущим программистам и маркетологам, для того чтобы быть конкурентоспособным на рынке трудовых ресурсов.

Список литературы

[1] Мария Решетникова. Искусственный Интеллект в Цифрах и Фактах [Электронный ресурс] – Режим доступа: <https://trends.rbc.ru/trends/industry/657963559a79474dd4bc9b88?from=copy>

[2] Понимание генеративного ИИ и трансформация бизнеса. [Электронный ресурс] – Режим доступа: https://info.idc.com/transform-genAI-ebook.html?utm_medium=idc-tagged&utm_source=idcweb-home&utm_campaign=us-idc-transform-genai-ebook&utm_idc=idh-0-idcus&_gl=1*bzzutu*_gcl_au*MTkyMzc3MTQzNi4xNzA3NDczNzg4*_ga*MTkzMDA2Mjg4OC4xNzA3NDczNzg4*_ga_Y7CNRMFF6J*MTcwNzQ3Mzc4OC4xLjAuMTcwNzQ3Mzc4OC42MC4wLjA.*_ga_541ENG1F9X*MTcwNzQ3Mzc4OC4xLjAuMTcwNzQ3Mzc4OC42MC4wLjA.

[3] Продажи и маркетинг [Электронный ресурс] – Режим доступа: https://store.hbr.org/search.php?search_query=Sales%20%26%20Marketing&ab=store_hp_nav_-_sales_marketing§ion=product

USING GENERATIVE ARTIFICIAL INTELLIGENCE IN BUSINESS PROCESS AUTOMATION

O.N.Shkor

Senior Lecturer

at the Department of Economics BSUIR

Abstract. The article discusses the prospects for the development of generative artificial intelligence and provides the opinion of analysts regarding the main trends in the use of artificial intelligence in various fields. Several potential applications for mobile applications with built-in AI are outlined. The possibilities of using generative artificial intelligence in programming, marketing, and the gaming industry are analyzed.

Keywords: generative artificial intelligence, marketing, gaming industry, chatbots, podcasts.

УДК 681.325

ПРОЦЕДУРА ВЫДЕЛЕНИЯ ПОДСИСТЕМ СВЯЗАННЫХ ФУНКЦИЙ ДЛЯ РЕШЕНИЯ ОПТИМИЗАЦИОННЫХ ЗАДАЧ ЛОГИЧЕСКОГО ПРОЕКТИРОВАНИЯ



Кириенко Н.А.

*Старший научный сотрудник Объединенного института проблем информатики НАН Б, кандидат технических наук, доцент
kir@newman.bas-net.by*

Н.А. Кириенко

Окончила Минский радиотехнический институт. Область научных интересов связана с автоматизацией процессов логического проектирования дискретных устройств, преобразованием и оптимизацией функциональных описаний логических схем.

Аннотация. Минимизация двухуровневых либо многоуровневых представлений булевых функций и систем является важнейшим этапом синтеза комбинационных схем, от которого зависят многие параметры логической схемы – площадь, быстродействие, энергопотребление. Повысить эффективность применения методов минимизации позволяет применение процедуры выделения подсистем связанных функций из системы булевых функций, описывающей поведение комбинационной схемы.

Рассматривается процедура выделения подсистем связанных функций, реализованная четырьмя алгоритмами. Приводится пример выполнения процедуры. Делается вывод об эффективности использования процедуры в системе логической оптимизации функциональных и структурных описаний логических схем FLC-2.

Ключевые слова: логическая схема, функциональное описание комбинационной схемы, системы булевых функций, подсистема связанных функций, мера связанности функций.

Введение. Актуальной задачей процесса автоматизированного проектирования дискретных устройств является технологически независимая оптимизация функциональных описаний цифровых блоков комбинационной логики заказных СБИС. Математическими моделями функционирования схем комбинационной логики являются системы булевых функций.

Одним из направлений логической оптимизации многоуровневых представлений систем булевых функций являются методы, основанные на выделении подсистем функций, которые имеют одинаковые части в областях определения функций выделенных подсистем. Такие подсистемы называются связанными. Связанность функций приводит к появлению большого числа одинаковых структурных частей (конъюнкций, алгебраических выражений, подфункций и др.) в оптимизированных формах представления функций, по которым строятся в дальнейшем комбинационные логические схемы. Чем сильнее связаны функции выделенной подсистемы, тем скорее можно ожидать, что в представлениях функций данной подсистемы будет больше одинаковых подвыражений и синтезированные логические схемы будут иметь меньшую сложность.

В настоящей работе предлагается описание алгоритма решения задачи выделения

подсистем связанных функций из многоуровневых представлений систем булевых функций. Для выделенных подсистем целесообразно проводить совместные минимизации в виде бинарных диаграмм решений (*BDD – Binary Decision Diagram*), названных *BDD – представлениями*, и в виде булевых сетей.

Постановка задачи. Задача выделения связанных по областям определения подсистем функций изучалась в [1], где были предложены различные меры связанности функций, в том числе и по общности формул разложения Шеннона для всех функций выделяемой подсистемы. Алгоритм позволял находить чаще всего только пары связанных функций. Модель для более эффективного алгоритма выделения связанных подсистем для совместной многоуровневой минимизации для двух видов исходных совместных формульных представлений систем функций предложена в [2].

Рассмотрим многоуровневое формульное представление системы булевых функций $F = \{f^1(x), \dots, f^m(x)\}$. Множество $R(F)$ внутренних формул, т.е. формул, задающих промежуточные (не выходные) булевы переменные для всех функций системы представлено выражением (1).

$$R(F) = R(f^1, \dots, f^m) = \bigcup_{i=1}^m R(f^i) \quad (1)$$

где $R(f^i)$ – множество внутренних формул в функции f^i .

Под мерой $q = M(F^m \cup \{f^{m+1}\})$ связанности системы функций $F^m \cup \{f^{m+1}\}$ будем понимать величину

$$q = M(F^m \cup \{f^{m+1}\}) = \frac{|R(f^1, \dots, f^m) \cap R(f^{m+1})|}{\max(|R(f^1, \dots, f^m)|, |R(\{f^{m+1}\})|)} \quad (2)$$

где через $|R|$ обозначена мощность множества R .

Величина меры связанности всегда находится в пределах отрезка $[0,1]$: от полного различия множеств формул исходного представления функций (либо подсистем функций) до их полного совпадения. Для выделенной подсистемы должно выполняться ограничение $q > ro$, где ro – некоторый заданный параметр для выполнения алгоритма.

Процедура выделения подсистем связанных функций из многоуровневых представлений систем булевых функций. Процедура построения подсистем связанных функций состоит из четырех алгоритмов. Алгоритм 1 является внешним алгоритмом внутри которого выполняются остальные три алгоритма.

Алгоритм 1.

Шаг 1. Выполнить разбиение системы на одновыходные блоки, в каждом блоке находится одна функция и все ее внутренние формулы, от которых она зависит. Определить n – количество блоков.

Шаг 2. Определить в каждом блоке список имен внутренних функций (или внутренних переменных). По этим спискам будет определяться связанность функций.

Шаг 3. Разбить исходную систему функций на подсистемы связанных функций согласно заданному критерию ro (выполнить алгоритм 2).

Алгоритм 2. Разбиение исходной системы функций S на подсистемы связанных функций согласно заданному параметру ro .

Исходные данные:

- набор одновыходных блоков двухуровневого описания, где каждый блок представляет одну функцию исходной системы;
- n – количество функций (блоков) в исходной системе;
- f_i – выходная функция i -го блока;

- f_j – выходная функция j -го блока;
- max – переменная, содержащая максимальную меру связанности для пары функций блока;
- ro – переменная, содержащая нижнее ограничение меры связности.
- rij – переменная, содержащая меру связанности функций f_i и f_j ;
- i_max, j_max – номера функций в паре, имеющей максимальную меру связанности;
- k – номер формируемой подсистемы;
- P_k – k -я подсистема;
- S – исходное состояние системы уравнений;
- Sk – текущее состояние системы уравнений (после удаления уравнений, подключаемых в подсистемы);
- pkm – мера связанности подсистемы P_k и функции f_m .

Результат: Двухуровневое описание из подсистем связанных функций.

Шаг 0. Текущее состояние системы $Sk = S$ (исходное состояние), $k = 1$.

Шаг 1. $i=1, max = 0$.

Шаг 2. $j=i+1$.

Шаг 3. Определить rij - меру связанности функций f_i и f_j (с помощью алгоритма 3).

Шаг 4. Если $rij > max, max = rij, i_max=i, j_max=j$.

Шаг 5. $j=j+1$.

Шаг 6. Если $j \leq n$ перейти на шаг 3.

Шаг 7. $i=i+1$.

Шаг 8. Если $i < n$ перейти на шаг 2.

Шаг 9. Формировать подсистему P_k из функций f_{i_max} и f_{j_max} .

Шаг 10. Удалить функции f_{i_max} и f_{j_max} из системы $Sk. n = n - 2$.

Шаг 10. Формировать список внутренних переменных подсистемы P_k .

Шаг 11. $m = 1, max = 0$.

Шаг 12. $m=m+1$.

Шаг 13. Определить pkm - меру связанности функций подсистемы P_k и f_m (с помощью алгоритма 4).

Шаг 14. Если $pkm > max, max = pkm, m_max = m$.

Шаг 15. $m = m + 1$.

Шаг 16. Если $m \leq n$ перейти на шаг 13.

Шаг 17. Если $pkm < ro$ перейти на 19.

Шаг 18. Добавить в подсистему P_k функцию с номером m_max . Удалить функцию с номером m_max из системы $Sk. n = n-1$. Перейти на 11.

Шаг 19. Формирование P_k завершено. Больше ни одной функции нельзя добавить без нарушения ограничения ($pkm \geq ro$).

Шаг 20. Если $n > 0, k = k + 1$, перейти на шаг 1.

Алгоритм 3 – определение меры связности rij функций f_i и f_j .

Определяем rij по формуле:

$$rij = S_mode/max_fij \quad (3)$$

где S_mode – количество одинаковых внутренних переменных, от которых зависят функции

f_i и f_j ;

max_fij – максимальное значение от числа внутренних переменных, от которых зависит функция f_i , и числа внутренних переменных, от которых зависит функция f_j .

Исходные данные:

– f_i – функция системы S_k ;

– f_j – функция системы S_k ;

Результат: мера связанности rij .

Шаг 1. Определить количества внутренних переменных в функциях f_i и f_j , и выбрать максимальное из них max_{fij} .

Шаг 2. Определить количество одинаковых внутренних переменных в функциях f_i и f_j – S_{mode} .

Шаг 3. Определить значение связанности функций rij по формуле 3.

Алгоритм 4 – определение меры связанности pkm подсистемы P_k и функции fm .

Определяем pkm по формуле:

$$pkm = S_{mode}/max_{pk}, \quad (4)$$

где S_{mode} – количество одинаковых внутренних переменных, от которых зависят подсистема P_k и функция fm ;

max_{pk} – максимальное значение от числа внутренних переменных, от которых зависит подсистема P_k , и числа внутренних переменных, от которых зависит функция fm .

Исходные данные:

P_k – подсистема связанных функций;

fm – функция системы S ;

Результат: мера связанности pkm подсистемы P_k и функции fm .

Шаг 1. Определить количества внутренних переменных в подсистеме P_k и в функции fm , и выбрать максимальное из них max_{pk} .

Шаг 2. Определить количество одинаковых внутренних переменных в подсистеме P_k и в функции fm – S_{mode} .

Шаг 3. Определить значение меры связанности pkm подсистемы P_k и функции fm по формуле 4.

Пример выполнения процедуры выделения подсистем связанных функций.

Процедура реализована в рамках системы логической оптимизации функциональных и структурных описаний логических схем $FLC-2$ [3]. Рассмотрим работу процедуры на примере, исходное задание которого представлено в листинге 1. Описания систем приводится на специализированном языке SF – структурно-функционального описания поведения комбинационной схемы [3]. Параметр $ro = 0,7$.

Листинг 1. Исходное задание системы функций.

```
TITLE z4
FORMAT SF
AUTHOR Lit
DATE 1-27-2000
PROJECT BenchMark
DCL_PIN
EXT
INP
x0 x1 x2 x3 x4 x5 x6
OUT
y0 y1 y2 y3
INTER
END_PIN
FUNCTION
LOG
7 4 0
y0=^x1*sf0+x1*sf1;
```

```

y1=^x1*sf2+x1*^sf2;
sf1=^x4*sf4+x4;
sf2=^x4*sf4+x4*^sf4;
sf0=x4*sf4;
sf4=^x2*sf10+x2*sf11;
y2=^x2*sf8+x2*^sf8;
sf11=^x5*sf12+x5;
sf8=^x5*sf12+x5*^sf12;
sf10=x5*sf12;
sf12=^x0*sf16+x0*sf17;
y3=^x0*sf18+x0*^sf18;
sf17=^x3*x6+x3;
sf18=^x3*x6+x3*^x6;
sf16=x3*x6;
END_LOG
END_FUNCTION
END_z4
    
```

Система содержит четыре функции: y_0 , y_1 , y_2 , y_3 ; и 11 внутренних переменных, от которых зависят функции. Вхождение внутренних переменных в каждую из функций (согласно шагу 1 алгоритма 1) представлено в таблице 1.

Таблица 1. Вхождение внутренних переменных в функции y_0 , y_1 , y_2 , y_3 .

y_0	y_1	y_2	y_3
$y_0 = ^x1 * sf0 + x1 * sf1$; $sf0 = x4 * sf4$; $sf1 = ^x4 * sf4 + x4$; $sf4 = ^x2 * sf10 + x2 * sf11$; $sf10 = x5 * sf12$; $sf11 = ^x5 * sf12 + x5$; $sf12 = ^x0 * sf16 + x0 * sf17$; $sf16 = x3 * x6$; $sf17 = ^x3 * x6 + x3$;	$y_1 = ^x1 * sf2 + x1 * ^sf2$; $sf2 = ^x4 * sf4 + x4 * ^sf4$; $sf4 = ^x2 * sf10 + x2 * sf11$; $sf10 = x5 * sf12$; $sf11 = ^x5 * sf12 + x5$; $sf12 = ^x0 * sf16 + x0 * sf17$; $sf16 = x3 * x6$; $sf17 = ^x3 * x6 + x3$;	$y_2 = ^x2 * sf8 + x2 * ^sf8$; $sf8 = ^x5 * sf12 + x5 * ^sf12$; $sf12 = ^x0 * sf16 + x0 * sf17$; $sf16 = x3 * x6$; $sf17 = ^x3 * x6 + x3$;	$y_3 = ^x0 * sf18 + x0 * ^sf18$; $sf18 = ^x3 * x6 + x3 * ^x6$;

Согласно шагу 3 алгоритма 2 определяем значение коэффициента связанности для всех пар функций с помощью алгоритма 3. Они представлены в таблице 2.

Таблица 2. Определение коэффициентов связанности для пар функций.

Пара функций	max_{fij}	S_{mode}	rij
$y_0 \ y_1$	8	6	$6 : 8 = 0,75$
$y_0 \ y_2$	8	3	$3 : 8 = 0,375$
$y_0 \ y_3$	8	0	0
$y_1 \ y_2$	7	3	$3 : 7 = 0,429$
$y_1 \ y_3$	7	0	0
$y_2 \ y_3$	4	0	0

Наибольший коэффициент связанности у пары $y_0 \ y_1$. Помещаем эту пару в подсистему P_1 (шаги 9, 10 алгоритма 2).

Листинг 2. Текущее состояние подсистемы P_1.

```

y0=^x1*sf0+x1*sf1;
sf0=x4*sf4;
sf1=^x4*sf4+x4;
sf4=^x2*sf10+x2*sf11;
sf10=x5*sf12;
sf11=^x5*sf12+x5;
sf12=^x0*sf16+x0*sf17;
sf16=x3*x6;
sf17=^x3*x6+x3;
y1=^x1*sf2+x1*^sf2;
sf2=^x4*sf4+x4*^sf4;
    
```

Согласно шагам 11 – 17 алгоритма 2 определяем значение коэффициента связанности между подсистемой P_1 и оставшимися функциями с помощью алгоритма 4. Они представлены в таблице 3.

Таблица 3. Определение коэффициентов связанности между подсистемой P_1 и оставшимися функциями.

Пара функций	<i>max_fij</i>	<i>S_mode</i>	<i>rij</i>
y0 y1, y2	9	3	3 : 9 = 0,333
y0 y1, y3	9	0	0

Наибольшее значение меры связанности у функций y0 y1, y2, но оно не удовлетворяет условию $rij > ro$. Поэтому построение подсистемы P_1 завершаем.

При построении системы P_2 выясняется, что оставшиеся функции y2 y3 являются не связанными. В результате, мы выделили одну подсистему связанных функций P_1, представленную в листинге 2.

Результаты экспериментального исследования. Представленная процедура реализована в рамках системы логической оптимизации функциональных и структурных описаний логических схем FLC-2 [3]. Экспериментальное исследование на серии описаний комбинационных логических схем представлено в [2].

Заключение. Процедура выделения подсистем связанных функций из многоуровневых представлений систем булевых функций позволила эффективно использовать различные методы оптимизации исходного описания комбинационных схем на этапе выполнения технологически независимой оптимизации. Эффект достигается за счет того, что каждая выделенная подсистема минимизируется на основе разложений Шеннона по своей (для каждой подсистемы) перестановке переменных разложения.

Список литературы

[1] Бибило П. Н., Позняк А. М. Выделение подсистем связанных функций из многоуровневого представления системы булевых функций // Информатика. 2020. Т. 17. № 1. С. 63 – 77.
 [2] Бибило, П.Н. Выделение из многоуровневого представления системы булевых функций подсистем для совместной логической минимизации / П.Н. Бибило, Н.А. Кириенко, В.И. Романов // Программные продукты и системы. 2023. Т. 36. № 4. С. 509 – 522. doi: 10.15827/0236-235X.142.
 [3] Бибило П.Н., Романов В.И. Система логической оптимизации функционально-структурных описаний цифровых устройств на основе продукционно-фреймовой модели представления знаний // Проблемы разработки перспективных микро- и нанoeлектронных систем. – 2020. Сб. трудов / под общ. ред. акад. РАН А.Л. Стемпковского. М.: ИППМ РАН. 2020. N 4. С. 9–16.

PROCEDURE FOR EXTRACTING SUBSYSTEMS OF CONNECTED FUNCTIONS FOR SOLVING OPTIMIZATION PROBLEMS OF LOGICAL DESIGN

N.A. Kirienko,

Senior Researcher of UIIP of NAS of Belarus,

PhD of Technical Sciences,

Associate Professor

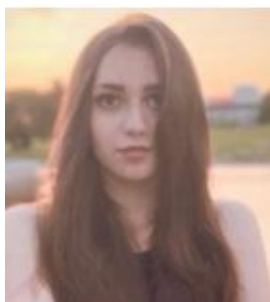
Abstract. Minimization of two-level or multi-level representations of Boolean functions and systems is the most important stage in the synthesis of combinational circuits, on which many parameters of a logical circuit depend - area, speed, power consumption. The efficiency of using minimization methods can be increased by using a procedure for extracting subsystems of connected functions from a system of Boolean functions that describes the behavior of a combinational circuit.

The procedure for extracting subsystems of connected functions, implemented by four algorithms, is considered. An example of performing the procedure is given. A conclusion is drawn about the effectiveness of using the procedure in the logical optimization system of functional and structural descriptions of logical circuits FLC-2.

Key words: logic circuit, functional description of the combinational circuit, systems of Boolean functions, the subsystem of connected functions, measure of function connectivity.

УДК 004.021:004.75

КЛАССИФИКАЦИЯ И АНАЛИЗ СНИМКОВ ЗЕМНОЙ ПОВЕРХНОСТИ



Е.И. Ковшер
Ассистент кафедры ЭВМ
БГУИР, магистрант
группы 356301
elizavetakovsher@mail.ru



Д.Ю. Перцев
Кандидат технических наук,
кафедра ЭВМ, БГУИР
pertsev@bsuir.by

Е.И. Ковшер

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием искусственного интеллекта и методов машинного обучения.

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. В статье выполнен анализ методов и алгоритмов классификации снимков земной поверхности. Исследованы подходы к эффективному разделению изображений на различные классы, что имеет важное значение для обработки больших объемов данных, получаемых с помощью спутниковых снимков и аэрофотосъемки.

Ключевые слова: Машинное обучение, нейронные сети, кластеризация, обработка геоинформационных данных.

Введение. В период стремительного технологического развития аэрофотосъемки и взрывного роста объема геоинформационных данных, алгоритмы автоматизированной классификации объектов на снимках земной поверхности становятся ключевым компонентом их обработки и анализа. Современные методы аэрофотосъемки, такие как спутниковые снимки, аэрофотосъемка с беспилотных летательных аппаратов (БПЛА), предоставляют огромные объемы данных, требующие эффективных методов обработки и классификации.

В современных условиях, когда географическая среда подвергается быстрым изменениям, задача классификации объектов на снимках земной поверхности становится более актуальной. Способность быстро и осознанно реагировать на эти изменения становится ключевой, а данные о землеустройстве становятся важными в таких областях, как градостроительство, сельское хозяйство и экология.

Геоинформационные системы (ГИС) играют важную роль в обработке и анализе снимков. ГИС обеспечивают интеграцию данных и их визуализацию, а передовые технологии в этой области значительно повышают эффективность классификации, предоставляя уникальные возможности для практического применения результатов [1].

Эффективные методы классификации объектов на снимках земной поверхности становятся ключевыми в различных областях, от оптимизации сельского хозяйства до

мониторинга городского развития и выявления изменений в экосистемах. Результаты классификации служат основой для стратегических решений.

В условиях постоянного увеличения объема данных обеспечение надежности и точности объектов на снимках земной поверхности становится критически важным. Это не только гарантирует корректность полученных данных, но и служит основой для принятия обоснованных решений в различных секторах человеческой деятельности [2].

Классическое обучение для классификации объектов на снимках земной поверхности. Первые алгоритмы основывались на применении теории математической статистики и появились в 1950-е годы. Целью алгоритма являлось формализовать процессы, связанные с выявлением статистических закономерностей в числовых данных, оценкой схожести точек в многомерных пространствах и определением направлений в данных. В настоящее время значительная часть современного интернета основывается на применении классических алгоритмов, которые, начиная с пятидесятых годов, продолжают эффективно решать различные формальные задачи, предоставляя механизмы для выявления структурных особенностей и обработки данных в различных областях. Методы классического обучения представлены на рисунке 1.



Рисунок 1. Методы классического обучения.

Обучение с учителем в контексте классификации снимков земной поверхности – это метод машинного обучения, основанный на использовании размеченных обучающих данных [3]. В этом подходе каждый снимок в обучающем наборе данных имеет соответствующую разметку, где каждый участок земли отнесен к определенному классу (категории). Например, к классам могут относиться такие объекты, как лес, вода, городская застройка и другие.

Классификация начинается с формирования обучающего набора данных, где для каждой области на снимке определен её класс. Этот набор используется для обучения модели машинного обучения. В качестве модели могут использоваться различные методы, включая метод опорных векторов (*SVM*), случайные леса или нейронные сети. Схема обучения представлена на рисунке 2.

При обучении модель извлекает признаки из снимков, такие как цвет, текстура, форма и другие характеристики, которые могут быть репрезентативными для определенных классов. Эти признаки используются для построения алгоритма, который может классифицировать новые, ранее не «виденные» данные.

После завершения обучения модель применяется к новым снимкам земной поверхности, где она предсказывает классы для каждой области. Этот процесс

обеспечивает автоматическую классификацию больших объемов данных, что может быть весьма полезным для различных целей, таких как мониторинг изменений в экосистемах, анализ использования сельскохозяйственных угодий или планирование городского развития.

Однако для успешного обучения модели требуется значительный объем размеченных данных, что может быть трудоемким и дорогостоящим процессом.

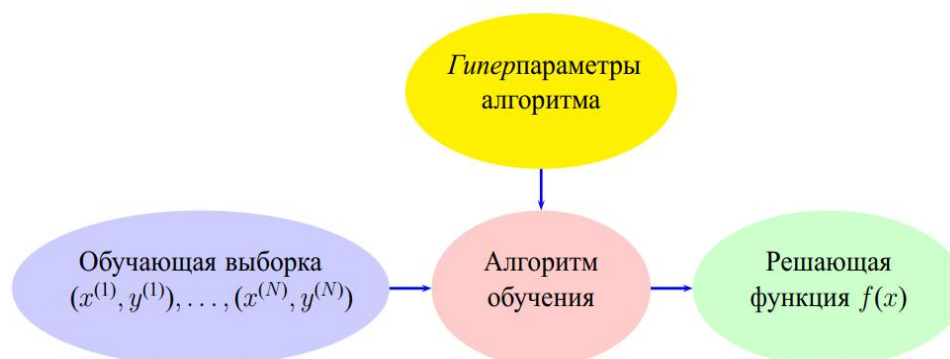


Рисунок 2. Схема обучения с учителем

Обучение без учителя в контексте анализа снимков земной поверхности представляет собой подход, призванный извлекать скрытые закономерности и структуры из больших объемов данных, не требуя предварительной разметки. Этот метод является эффективным инструментом в условиях, когда получение размеченных обучающих данных может быть сложным или дорогостоящим [3].

Одним из ключевых подходов в обучении без учителя для анализа снимков земли является кластеризация. Данный метод направлен на группировку схожих участков земли на основе их признаков, таких как цвет, текстура или другие характеристики. Например, алгоритмы кластеризации могут объединять участки с похожими спектральными характеристиками, что может указывать на схожие типы покрытия, такие как лес, вода или городская застройка.

На первом этапе выбирается метод кластеризации, например, k -средних или иерархическая кластеризация, и определяются признаки для анализа, такие как спектральные и текстурные характеристики снимков. Далее данные подвергаются предобработке, включая нормализацию и устранение шума. Применяется выбранный метод кластеризации, который группирует схожие участки земли в кластеры. Результаты оцениваются с использованием таких метрик, как коэффициент силуэта, затем выполняется интерпретация каждого кластера, связывая его с определенным типом покрытия (вода, лес, городская застройка и т. д.).

Обучение включает в себя итерации для уточнения параметров и выбора оптимальных методов предобработки, а также использование метрик для объективной оценки результатов. Этот подход позволяет автоматически выделять структуры на земной поверхности без явной разметки данных, что может быть полезным в мониторинге изменений в экосистемах, анализе использования сельскохозяйственных угодий и планировании городского развития.

Другой важный аспект обучения без учителя в данной области – это методы понижения размерности. Эти методы позволяют сократить число признаков, оставляя наиболее информативные аспекты данных [4]. Это может быть полезным при работе с большими массивами информации, такими как космические снимки, где эффективное извлечение ключевых характеристик спектральных данных может улучшить процесс анализа.

Примерами методов обучения без учителя для анализа снимков земли также являются методы ассоциативного правила, позволяющие выявлять «интересные» взаимосвязи и закономерности между различными областями земной поверхности, а также методы генеративных моделей, которые строят вероятностные модели для данных, что может быть полезным при синтезе новых областей, которые не анализировались ранее.

Таким образом, обучение без учителя в анализе снимков земной поверхности предоставляет мощный инструмент для обнаружения паттернов, выделения характеристик и сжатия информации, что может быть ключевым в условиях стремительного развития технологий аэросъемки и обработки геоинформационных данных.

Ансамбли для классификации снимков земной поверхности. Ансамбли в контексте классификации снимков земной поверхности представляют собой метод машинного обучения, основанный на комбинировании результатов нескольких моделей для улучшения общей производительности и точности предсказаний.

При построении ансамбля создаются различные модели, используя вариации параметров. Эти модели затем объединяются для формирования более надежного и точного прогноза. В контексте классификации снимков земной поверхности, ансамбли могут быть применены для улучшения определения типов покрытия, таких как вода, лес, городская застройка и другие. Методы ансамблей представлены на рисунке 3.



Рисунок 3. Методы ансамблей

Популярные методы ансамблей включают в себя случайный лес (англ. *Random Forest*) и бустинг (англ. *Boosting*). Случайный лес – это алгоритм, который строит несколько деревьев решений при обучении и объединяет их для формирования более надежного и устойчивого предсказания. Каждое дерево строится на основе случайной подвыборки данных и случайного подмножества признаков. Это обеспечивает разнообразие моделей в ансамбле, что способствует улучшению обобщающей способности. При классификации каждое дерево голосует за определенный класс, и класс с наибольшим количеством голосов становится окончательным предсказанием. Бустинг – это метод, который последовательно обучает модели, уделяя больше внимания неправильно классифицированным данным на каждом этапе. Это позволяет улучшать точность предсказаний по мере итераций. Один из популярных вариантов бустинга – градиентный бустинг (англ. *Gradient Boosting*), который минимизирует ошибку, используя градиент функции потерь. Этот процесс приводит к созданию ансамбля, где каждая модель корректирует ошибки предыдущей, что приводит к повышению общей производительности [5].

Применение ансамблей в анализе снимков земной поверхности позволяет значительно улучшить обобщающую способность моделей и повысить эффективность классификации при обработке обширных объемов геоинформационных данных. Этот

подход становится важным элементом в арсенале методов для обеспечения точности и стабильности в решении задач классификации снимков земной поверхности.

Нейросети и глубокое обучение для классификации снимков земной поверхности. Глубокое обучение и нейросети представляют современный и мощный подход к классификации снимков земной поверхности. Нейросети, особенно свёрточные (*CNN*), являются ключевыми инструментами, способными автоматически извлекать сложные паттерны и признаки из данных. Методы глубокого обучения представлены на рисунке 4. Процесс начинается с тщательной подготовки размеченных данных, где каждый снимок земли связан с соответствующим классом.

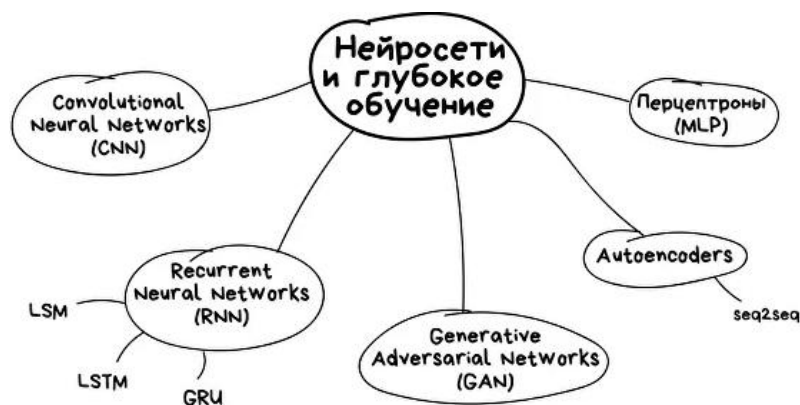


Рисунок 4. Методы глубокого обучения

Выбор подходящей архитектуры нейросети играет важную роль в обработке изображений. Свёрточные нейронные сети оптимизированы для эффективного извлечения признаков, что особенно важно при анализе снимков земли. Обучение модели включает в себя передачу данных через слои сети и корректировку весов на основе обучающих данных.

После завершения обучения модель проверяется на проверочном наборе данных для оценки её обобщающей способности. Затем выполняется тестирование на отдельном тестовом наборе для проверки точности и эффективности на новых данных. Тонкая настройка гиперпараметров модели требуется для предотвращения переобучения и оптимальной производительности [6].

В дальнейшем обученная модель может быть успешно применена к новым снимкам земной поверхности, автоматически классифицируя объекты в соответствии с предварительно определенными классами. Важно отметить, что для успешного применения глубокого обучения требуется значительный объем размеченных данных и вычислительные ресурсы.

Рекуррентные нейронные сети (англ. *RNN*) также могут быть включены в набор методов для анализа снимков земной поверхности. В отличие от сверточных нейронных сетей, *RNN* специализируются на обработке последовательных данных и могут быть полезны при работе с временными рядами или изображениями, где важна пространственная и временная зависимость.

В контексте анализа снимков земли, *RNN* могут использоваться, например, для учёта изменений в природных условиях или сезонных колебаний. Модели *RNN* позволяют анализировать последовательности снимков с течением времени и выявлять паттерны, которые могут быть важны для понимания динамики на земной поверхности.

Обучение *RNN* также включает в себя передачу данных через временные слои, и адаптацию весов в соответствии с обучающими данными. Такие модели предоставляют возможность учесть контекст и историю данных при классификации [7].

Однако, следует отметить, что *RNN* также могут сталкиваться с проблемой исчезающего градиента, особенно при работе с длинными последовательностями данных. В таких случаях, более современные архитектуры, такие как долгая краткосрочная память (англ. *LSTM*) или *Gated Recurrent Unit (GRU)*, могут использоваться для решения указанной проблемы.

Таким образом, *RNN* предоставляют дополнительные возможности для анализа снимков Земли, особенно там, где важна временная динамика и изменения с течением времени.

Генеративные состязательные сети (англ. *GAN*) представляют собой другой метод, состоящий из двух основных компонентов: генератора и дискриминатора.

Генератор создает новые изображения, имитируя реальные данные. В контексте снимков земной поверхности, генератор может создавать синтетические изображения, которые соответствуют различным классам покрытия, таким как леса, водные поверхности, города и т.д. Дискриминатор, с другой стороны, оценивает, насколько изображение соотносится между реальным и сгенерированным.

Процесс обучения представляет собой взаимодействие между генератором и дискриминатором. Генератор стремится к совершенствованию своей способности создания реалистичных изображений, в то время как дискриминатор улучшает свою способность отличать поддельные изображения от настоящих [8].

В контексте снимков земной поверхности, *GAN* могут использоваться для создания новых, несуществующих изображений, которые, тем не менее, выглядят правдоподобно. Это может быть полезным в ситуациях, где не хватает данных или для аугментации существующего набора данных.

Однако стоит отметить, что обучение *GAN* может быть сложным, и требует тщательной настройки параметров для достижения стабильного и качественного результата.

Многослойные перцептроны (англ. *MLP*) представляют собой класс нейронных сетей, состоящих из нескольких слоев: входного слоя, скрытых слоев и выходного слоя. В контексте классификации снимков земной поверхности, *MLP* являются эффективным инструментом для извлечения сложных закономерностей из данных:

- входной слой. На этом слое каждый нейрон представляет собой определенный признак снимка, такой как цвет, текстура или другие характеристики.

- скрытые слои. Эти слои служат для извлечения абстрактных иерархических признаков из входных данных. Каждый нейрон в скрытом слое соединен со всеми нейронами предыдущего и последующего слоя.

- выходной слой. На этом слое каждый нейрон представляет собой класс объекта или тип покрытия земли (например, лес, вода, город). Выходы этих нейронов могут интерпретироваться как вероятности принадлежности объекта к каждому классу.

Процесс обучения включает в себя передачу данных через сеть, расчет предсказаний и сравнение их с фактическими метками классов. Алгоритм обратного распространения ошибки используется для корректировки весов нейронов, чтобы минимизировать ошибку предсказания. Этот процесс повторяется на множестве эпох обучения [9].

Преимущества *MLP* включают их способность моделировать сложные нелинейные отношения между признаками и классами. Они также могут автоматически извлекать признаки из данных, что упрощает процесс обучения.

Однако для успешного обучения *MLP* требуется большой объем размеченных данных, и процесс обучения может быть вычислительно интенсивным, особенно при использовании глубоких архитектур.

Заключение. Методы классификации снимков земной поверхности представляют разнообразные подходы, варьирующиеся от классических алгоритмов до передовых

технологий вроде нейросетей и глубокого обучения. Каждый из рассмотренных методов обладает своими преимуществами и ограничениями. Выбор метода машинного обучения для классификации и анализа снимков земной поверхности зависит от различных факторов, включая характеристики данных, объем данных, требования к точности, а также доступные ресурсы. Характеристики данных, такие как разрешение изображений, спектральные характеристики и текстурные особенности, могут влиять на эффективность различных методов. Объем данных, в свою очередь, может потребовать использование вычислительно эффективных алгоритмов или параллельных вычислений для обеспечения быстрой обработки. В данной области часто используется комбинация различных методов для достижения наилучших результатов. В дальнейшем планируется практическая реализация представленных методов классификации и сопоставление полученных результатов между собой.

Список литературы

- [1] Капралов Е.Г., Кошкарев А.В., Тикун В.С. и др. Основы геоинформатики. Уч. пособие. – М.: Изд. центр «Академия», 2004. – 480 с.
- [2] Шовенгердт Р.А. Дистанционное зондирование. Модели и методы обработки изображений. М.: Техносфера, 2010. – 560 с.
- [3] Васенков Д.В. Методы обучения искусственных нейронных сетей // Компьютерные инструменты в образовании. 2007. № 1. – 20 с.
- [4] Хайкин С. Нейронные сети. Полный курс. – М.: Вильямс, 2006. – 1104 с.
- [5] Viola P., Jones M.J. Robust real-time face detection // International Journal of Computer Vision, 2004 – 137 с.
- [6] Татьянkin, В. М. Подход к формированию архитектуры нейронной сети для распознавания образов / В. М. Татьянkin // Вестник Югорского государственного университета, 2016, – 61 с.
- [7] Никитин, К.В. Новый подход к применению рекуррентных нейронных сетей для решения задач распознавания образов, 2007 – 85 с.
- [8] Гудфеллоу, И.Дж. Генеративные состязательные сети в системах обработки нейронной информации / И.Дж. Гудфеллоу, Дж. Пуже-Абади, М. Мирза, Б. Сюй, Д. Вард-Фарли, С. Озер, А.С. Курвиль, Ю. Бенжио // 27-я Ежегодная конференция по нейронным системам. – 2014. – С. 2672-2680.
- [9] Толмачев С.Г., Системы искусственного интеллекта. Нейросетевые модели: учебное пособие, Балт. гос. техн. ун-т. – СПб., 2011 г. – 170 с.

Авторский вклад

Ковшер Елизавета Игоревна – обзор методов и алгоритмов, применяемых для классификации снимков земной поверхности.

Перцев Дмитрий Юрьевич – постановка задачи исследования, анализ полученных результатов.

CLASSIFICATION AND ANALYSIS OF EARTH SURFACE IMAGES

E. KOVSHER

*Assistant at the Department of
Electronic Computing Machines,
BSUIR, Master's student group
356301*

D. PERTSAU

*PhD,
Electronic Computing Machines
Department, BSUIR*

Abstract. In this article, an analysis of methods and algorithms for the classification of images of the Earth's surface has been conducted. Approaches to the effective separation of images into various classes were explored, which is crucial for processing large volumes of data obtained from satellite imagery and aerial photography.

Keywords: Machine Learning, Neural Networks, Clustering, Geoinformation Data Processing.

УДК 37.018: [330:004+331.522] (470+571)

ЭЛЕКТРОННОЕ ОБУЧЕНИЕ ДЛЯ РАЗВИТИЯ КАДРОВОГО ПОТЕНЦИАЛА: ОПЫТ РОССИЙСКОЙ ФЕДЕРАЦИИ¹



И.П. Сидорчук

Заместитель директора по научно-методической работе Института информационных технологий БГУИР, кандидат юридических наук, доцент
irina_sidorchuk@bsuir.by



А.А. Охрименко

Директор Института информационных технологий БГУИР, кандидат технических наук, доцент
ohrimenko@bsuir.by



Е.Г. Крысь

Заведующий сектором учебно-методического отдела Института информационных технологий БГУИР, магистр управления и права
kryсь_eg@bsuir.by

И.П. Сидорчук

Окончила Белорусский государственный университет и аспирантуру в Национальной академии наук Беларуси. Сфера научных интересов: исследования в области нормотворческой деятельности, государственного управления, кадровой политики, оценки регулирующего воздействия, устойчивого развития общества, непрерывного образования и профессионального развития руководителей и специалистов в условиях цифровой экономики.

А.А. Охрименко

Окончил Минское высшее инженерное зенитное ракетное училище противовоздушной обороны и адъюнктуру в Военной инженерной радиотехнической академии противовоздушной обороны им. Л.А. Говорова. Сфера научных интересов: исследования в области цифровой трансформации и электронного правительства, информационно-коммуникационных технологий, вопросы устойчивого развития общества, государственного управления, непрерывного профессионального образования.

Е.Г. Крысь

Окончила Белорусский институт правоведения, Институт государственной службы и магистратуру Академии управления при Президенте Республики Беларусь. Аспирант Национального центра законодательства и правовых исследований Республики Беларусь. Сфера научных интересов: исследования в области государственного управления, оценки регулирующего воздействия, непрерывного образования и профессионального развития руководителей и специалистов в условиях цифровой экономики, устойчивого развития общества.

¹ Настоящая статья подготовлена по результатам выполнения НИР «Разработка модели» цифровых компетенций и «Концепции подготовки образовательного контента для обучения работников государственных органов и организаций по вопросам цифрового развития». Выполнение вышеуказанной НИР осуществлялась авторами статьи в рамках реализации этапа 1 «Разработка образовательного контента для обучения работников государственных органов и организаций по вопросам цифрового развития» по мероприятию 2 «Разработка образовательного контента для курсов повышения квалификации работников государственных органов и организаций по вопросам цифрового развития» подпрограммы 1 «Информационно-аналитическое организационно-техническое сопровождение цифрового развития» Государственной программы «Цифровое развитие Беларуси на 2021-2025 годы».

Аннотация. В статье анализируется опыт Российской Федерации по вопросам организации обучения руководителей и специалистов для работы в условиях цифровой экономики. Внимание акцентируется на онлайн возможностях современного образования. Делается вывод о том, что многие образовательные площадки, цифровые университеты внедряют в практику электронное обучение. Растет востребованность дистанционных технологий.

Ключевые слова: цифровая трансформация, цифровая экономика, развитие цифровых компетенций, кадры для цифровой экономики, дополнительное профессиональное развитие.

В современных условиях в Беларуси, как и в других странах, приоритетным направлением является подготовка кадров для работы в условиях цифровой экономики. В государственной программе «Цифровое развитие Беларуси» на 2021–2025 годы, утвержденной постановлением Совета Министров Республики Беларусь от 2 февраля 2021 г. № 66 (далее – Программа цифрового развития), определено, что необходимо обеспечить доступность образования, основанного на применении современных информационных технологий как для повышения качества образовательного процесса, так и для подготовки граждан к жизни и работе в условиях цифровой экономики. Указанный документ подтверждает актуальность непрерывного профессионального развития в ИТ-сфере. В соответствии с показателями Программы цифрового развития доля специалистов, ответственных за вопросы информатизации в государственных органах и организациях, прошедших обучение в сфере цифрового развития к 2025 г., должна составить 40 %. Кроме того, предусматривается создание образовательной платформы повышения «цифровой грамотности» населения. Все это требует конкретных действий со стороны учреждений высшего образования, иных заинтересованных в части разработки и реализации новых образовательных продуктов по вопросам цифрового развития. Перспектива развития образования состоит в постепенном обучении граждан, а также руководителей и работников государственных организаций цифровым навыкам с целью увеличения численности кадров, обладающих новыми компетенциями, «цифровым», критическим, креативным мышлением.

В этой связи актуально проанализировать подходы зарубежных стран, в частности, Российской Федерации, в которой достигнуты значимые результаты в части обучения цифровым навыкам, а также выработать предложения и рекомендации по формированию цифровых компетенций у руководителей и специалистов в Беларуси, определив виды образовательных программ, актуальные технологии и обосновав преимущества электронного обучения.

В Российской Федерации в рамках реализации указов Президента Российской Федерации от 7 мая 2018 г. № 204 «О национальных целях и стратегических задачах развития Российской Федерации на период до 2024 года» и от 21.07.2020 г. № 474 «О национальных целях развития Российской Федерации на период до 2030 года», в том числе с целью решения задачи по обеспечению ускоренного внедрения цифровых технологий в экономике и социальной сфере, Правительством Российской Федерации сформирована национальная программа «Цифровая экономика Российской Федерации». Эта программа утверждена протоколом заседания президиума Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам от 4 июня 2019 г. № 7, так, к 2024 году планируется обучить 50 тыс. государственных и муниципальных служащих цифровым навыкам.

24 декабря 2018 г. утвержден Паспорт национальной программы «Цифровая экономика Российской Федерации» (далее – Паспорт) [1]. Основной целью этого документа является создание условий для формирования рынка труда, обеспеченного квалифицированными и конкурентоспособными кадрами для цифровой экономики через трансформацию всех уровней образования, внедрения программ переобучения в

компаниях и ведомствах. Составной частью Паспорта является Федеральный проект «Кадры для цифровой экономики». В нем установлены следующие целевые индикаторы:

«270 тыс. работающих специалистов, включая руководителей организаций и представителей органов исполнительной власти, прошли обучение по компетенциям цифровой экономики (начиная с 2019 года до конца 2024 года);

10 млн человек прошли обучение по онлайн программам развития цифровой грамотности (начиная с 2019 года до конца 2024 года);

120 тыс. человек принято на программы высшего образования в сфере информационных технологий до 31 октября 2024 года;

созданы и функционируют 50 центров ускоренной подготовки специалистов совместно с компаниями цифровой экономики (начиная с 2020 года по 31 декабря 2023 года)» [2].

В феврале 2019 года создан Центр подготовки руководителей цифровой трансформации (далее – Центр подготовки) (<https://cdto.ranepa.ru/>) – самая массовая организация по обучению сотрудников государственного сектора. Центр подготовки является подразделением Высшей школы государственного управления РАНХиГС. Основными направлениями деятельности Центра подготовки являются:

– проведение образовательных программ для государственных служащих (обучение «цифре»);

– карьерное развитие специалистов цифровой трансформации;

– реализация мониторинга ИИ-зрелости² органов госуправления.

Центр подготовки разработал совместно с Федеральным институтом развития образования и реализует следующие программы повышения квалификации, направленные на:

1 *развитие профессионально-технических компетенций* – «Цифровая трансформация: от первых шагов до максимальных результатов»; «Цифровая стратегия: разработка и практическая реализация»; «Введение в государственное и муниципальное управление»; «Управление цифровым развитием: как достичь максимального результата»; «Цифровое развитие: организационные изменения»; «Управление данными: современные подходы»; «Практика управления ИТ-инфраструктурой»; «Преподавание онлайн: от первого вебинара к профессиональной модерации»;

2 *развитие личностно-управленческих компетенций* – «Целеполагание и принятие решений»; «Как найти хорошую работу: руководство от практикующего эксперта»; «Клиентоцентричное государство: культура, правила и инструменты»; «Создание и развитие высокоэффективной команды в условиях ТАСИ – мира».

В структуре Центра подготовки в апреле 2020 г. создан Центр оценки и кадрового развития специалистов в области цифровой трансформации (Рекрутинговый центр Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации).

Ключевая цель его создания – оказание квалифицированной помощи подразделениям цифровой трансформации государственных органов власти и их подведомственным организациям в поиске и оценке квалифицированных, перспективных и мотивированных к профессиональному развитию кадров в IT сфере и цифровой трансформации с 100% закрытием вакансий (рисунок 1).

² ИИ-зрелость – это международный опыт систематизации практик по внедрению Data Science-подходов в бизнес-процессы



Рисунок 1. Ключевые направления деятельности Центра подготовки руководителей цифровой трансформации (<https://cdto.ranepa.ru/>)

В соответствии с действующим законодательством Российской Федерации профессиональное развитие гражданского служащего осуществляется в течение всего периода прохождения им службы. Оно включает в себя не только дополнительное профобразование, семинары, тренинги, мастер-классы, конференции, круглые столы, служебные стажировки, но и самостоятельное изучение государственными служащим учебных материалов по соответствующей его служебной деятельности тематике, а также дистанционные образовательные курсы (пункты 4, 6 Положения о порядке осуществления профессионального развития государственных гражданских служащих Российской Федерации, утвержденного Указом Президента Российской Федерации от 21 февраля 2019 г. № 68).

При этом Министерство труда и социальной защиты Российской Федерации уполномочено по согласованию с Администрацией Президента Российской Федерации и Аппаратом Правительства Российской Федерации ежегодно утверждать приоритетные направления профессионального развития федеральных государственных гражданских служащих (пункт 22 вышеуказанного положения). Так, например, приоритетные направления на период до 2024 г. формировались исходя из национальных целей развития Российской Федерации на период до 2024 года. В их основе положения Указа Президента Российской Федерации от 7 мая 2018 года № 204 «О национальных целях и стратегических задачах развития Российской Федерации на период до 2024 года», национальные цели развития Российской Федерации на период до 2030 года, определенные Указом Президента Российской Федерации от 21 июля 2020 года № 474 «О национальных целях развития Российской Федерации на период до 2030 года», а также основные направления деятельности Правительства Российской Федерации на период до 2024 года, утвержденные Правительством Российской Федерации 29 сентября 2018 года № 8028 п-П13.

Кроме этого, Федеральным законом от 27 июля 2004 года № 79-ФЗ «О государственной гражданской службе Российской Федерации», Указом Президента РФ от 28 декабря 2006 г. № 1474 «О дополнительном профессиональном образовании государственных гражданских служащих Российской Федерации», постановлением Правительства Российской Федерации от 06.05.2008 № 362 «Об утверждении государственных требований к профессиональной подготовке, повышению квалификации и стажировке государственных гражданских служащих Российской Федерации» предусмотрена разработка индивидуальных планов профессионального развития государственного гражданского служащего. Индивидуальный план разрабатывается им в соответствии с должностным регламентом совместно с непосредственным руководителем сроком на три года и включает в себя перечень мероприятий по дополнительному профессиональному образованию государственных гражданских служащих

(профессиональная переподготовка, повышение квалификации). Эти мероприятия позволяют приобрести, углубить и обновить в процессе обучения профессиональные знания, умения и навыки для обеспечения соответствия уровня профессиональной подготовки гражданских служащих содержанию и объему должностных обязанностей по замещаемой должности.

Министерство труда и социальной защиты Российской Федерации в целях создания и обеспечения функционирования на гражданской службе системы квалификационных требований, отвечающих новым реалиям, и в рамках практической реализации положений части восьмой статьи 12 Федерального закона от 27 июля 2004 г. № 79-ФЗ «О государственной гражданской службе Российской Федерации» разработало Справочник квалификационных требований к специальностям, направлениям подготовки, знаниям и умениям, которые необходимы для замещения должностей государственной гражданской службы с учетом области и вида профессиональной служебной деятельности государственных гражданских служащих. Этот документ определяет виды цифровых навыков и содержание знаний.

Методические пояснения к вышеназванному документу дают определение квалификации как определенного уровня компетенции, которая характеризует готовность выполнять профессиональную деятельность. Так, к базовой группе компетенций относятся знания и навыки вне занимаемой должности, а ко второй – по определенной специальности. В числе базовых установок для всех должностей гражданской службы применяется правило по владению навыками в области информационно-коммуникационных технологий: общие знания информационных технологий и использования персонального компьютера; общие навыки работы на персональном компьютере; основы защиты информации и информационная безопасность; общие навыки работы в сети «Интернет»; знания основных офисных программ.

Во исполнение абзаца четвертого подпункта «б» пункта 4 Указа Президента Российской Федерации от 21 февраля 2019 г. № 68 «О профессиональном развитии государственных гражданских служащих Российской Федерации» на базе федеральной государственной информационной системы «Единая информационная система управления кадровым составом государственной гражданской службы Российской Федерации» (<https://gossluzhba.gov.ru>) создан Единый специализированный информационный ресурс (далее – Единый ресурс).

В соответствии с пунктом 3 Положения «О едином специализированном информационном ресурсе, предназначенном для профессионального развития государственных гражданских служащих Российской Федерации», утвержденного постановлением Правительства Российской Федерации от 15 августа 2019 г. № 1056 Единый ресурс создается в целях широкого применения информационно-коммуникационных технологий в ходе планирования, организации и реализации мероприятий по профессиональному развитию гражданских служащих, которые реализуются организациями, осуществляющими образовательную деятельность по дополнительным профессиональным программам для гражданских служащих, и (или) иными организациями, а также государственными органами. Сервис позволит гражданским служащим сформировать индивидуальную образовательную траекторию на основе цифрового профиля. Доступ к Единому ресурсу является бесплатным³.

У гражданских служащих появятся «личные кабинеты», через которые можно будет:
–оценить свои знания и умения;

³ Дата окончательного ввода в действие ресурса не определена. На момент подготовки статьи доступ к Единому ресурсу был ограничен.

- сформировать направления профессионального развития;
- получить доступ к обучению по дополнительным профессиональным программам;
- самостоятельно изучить образовательные материалы.

Согласно информации Министерства труда и социальной защиты Российской Федерации, внедрение единого ресурса обеспечит получение гражданами служащими знаний и навыков в непрерывном формате в соответствии с современными подходами к развитию персонала. В рамках этого ресурса появятся возможности оценить свои знания и умения, получить информацию о приоритетных направлениях профессионального развития от образовательных организаций и государственных органов, обучиться по дополнительным профессиональным программам и др. [3].

Основными функциями Единого ресурса являются:

- а) ознакомление гражданских служащих с информацией о приоритетных направлениях профессионального развития гражданских служащих и о мероприятиях по профессиональному развитию;
- б) ознакомление гражданских служащих с информацией о дополнительных профессиональных программах для гражданских служащих, реализуемых на основании государственного образовательного сертификата на дополнительное профессиональное образование государственного гражданского служащего Российской Федерации образовательными организациями, включенными в реестр исполнителей государственной услуги по реализации дополнительных профессиональных программ для государственных гражданских служащих Российской Федерации на основании образовательных сертификатов;
- в) проведение гражданскими служащими самостоятельной оценки своих знаний и умений;
- г) формирование в автоматизированном виде направлений профессионального развития гражданских служащих;
- д) участие гражданских служащих в мероприятиях по профессиональному развитию, реализуемых в том числе с использованием информационно-коммуникационных технологий;
- е) доступ гражданских служащих к информационным ресурсам;
- ж) самостоятельное изучение гражданскими служащими актуальных обучающих, методических, аналитических и иных материалов образовательного характера по тематике, соответствующей направлению профессиональной служебной деятельности;
- з) осуществление оценки качества реализации мероприятий по профессиональному развитию в том числе посредством заполнения гражданскими служащими и сотрудниками кадровых подразделений государственных органов формы опроса в электронном виде;
- и) учет мероприятий по профессиональному развитию и их результатов в электронном личном деле;
- к) формирование статистических и аналитических отчетных материалов.

Выше отмечалось, что в Беларуси создается образовательная платформа повышения «цифровой грамотности» населения в этой связи с учетом отмеченного опыта РФ актуально на нормативном уровне определить функции этой платформы, порядок подготовки и размещения образовательных программ для профессионального развития, а также подходы к оценке полученных знаний и т.д.

Профессиональное развитие цифровых компетенций государственных служащих требует построения надежной основы, обеспечивающей качественное обучение: необходимо определить актуальные для всех категорий персонала цифровые компетенции, создать правовую базу, построить сотрудничество с поставщиками образовательных услуг и технологий, а также на основе программ непрерывного обучения и организации проводить мониторинг образовательного процесса и его результатов.

Учреждения высшего образования для решения задач в части подготовки кадров для цифровой экономики должны сами трансформироваться в цифровые университеты. В цифровых учреждениях высшего образования «будут задействованы различные инновационные возможности цифровых образовательных платформ вузов: персональные образовательные траектории с применением искусственного интеллекта, «точки кипения», фабрики пилотирования, проектно-образовательные интенсивы, форматы цифрового следа. Усилятся требования к цифровой академической мобильности студентов и преподавателей [4]». Например, Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации совместно с университетом «Иннополис» создали консорциум учебных заведений по подготовке кадров для цифровой экономики. Сейчас в составе консорциума числится 185 организаций высшего и 101 организация среднего профессионального образования по всей России. Консорциум станет площадкой, которая объединит образовательные организации в работе над актуализацией и стандартизацией образовательных программ с учетом потребностей современного рынка в условиях цифровой экономики. Будут созданы модули обучения для преподавателей, методистов и студентов, в разработке которых помогут эксперты ИТ-индустрии.

Высшая школа государственного управления РАНХиГС в марте 2022 года запустила проект – глобальный цифровой университет «**Университет 2035**» [5] как основная образовательная платформа по развитию навыков в сфере ИКТ (<https://www.2035.university/>). Университет 20.35 – это 1 312 000 пользователей, 150 000 единиц контента (образовательных онлайн-курсов, лекций и статей доступны на платформе). 257 000 человек прошли обучение на платформе за время работы университета. Каждый человек на платформе Университета 20.35 может получить доступ к обучающим курсам. Для начала обучения необходимо пройти регистрацию и выбрать курсы в личном кабинете. Цифровая диагностика, которую проходит пользователь, и накопление данных об образовательной траектории дают возможность платформе давать более точные рекомендации для развития каждого обучающегося. Обучение – поэтапное, с постоянной обратной связью и выбором ближайшего шага развития. Результаты каждого шага пользователь может увидеть в своем профиле компетенций в личном кабинете, оценить для себя правильность пути и, при поддержке рекомендательных систем, сделать следующий шаг.

На платформе размещаются сервисы, обеспечивающие профессиональное развитие человека в цифровой экономике. Например, сервис «**Готов к цифре**» – это проект о безопасном и эффективном использовании цифровых технологий для людей самых разных уровней цифровых компетенций (<https://xn--blabhljwatnyu.xn--plai/>). На данном ресурсе собраны самые разные инструменты оценки и развития цифровой грамотности и компетенций цифровой экономики, которые помогут определить слабые стороны и покажут вектор для дальнейшего развития цифровых компетенций. Здесь размещены 150 тестовых заданий, позволяющих определить цифровой уровень развития. Сервис диагностики цифровых навыков объединяет самые разные инструменты оценки. После прохождения процедуры авторизации и тестирования, будет определен вектор для дальнейшего развития или совершенствования цифровых компетенций. После чего из 300 предложенных экспертами видео (видеоуроков), учебных пособий, инструкций, статей, нормативных документов можно выбрать и изучить необходимый курс по цифровым продуктам и новейшим технологиям для повышения цифровой грамотности сотрудников или привлечения специалистов нецифровых профессий в новую область профессиональных знаний.

Образовательный контент сгруппирован следующим образом: искусственный интеллект; анализ больших данных; онлайн и смешанное обучение; *Python*; 3D моделирование; цифровые навыки; анализ потребностей бизнеса; педагогический дизайн;

персонализированное образование; платформы онлайн образования. Поиск необходимой информации на сервисе осуществляется по ключевым словам или применяется фильтр по актуальным цифровым компетенциям, например, таким как: управление информацией и данными; коммуникации в цифровой среде; безопасность и защита данных; создание цифрового контента; решение проблем в «цифре»; саморазвитие в условиях неопределенности; креативное мышление; критическое мышление.

На сервисе «Готов к цифре» размещен ресурс «Удаленная занятость», который создан при участии компаний-участников Консорциума компаний по развитию цифровой грамотности в рамках задачи по развитию цифровой грамотности населения федерального проекта «Кадры для цифровой экономики» национальной программы «Цифровая экономика РФ». Здесь собрано более 120 лучших практик, сервисов и инструкций от Ростелекома, Сбербанка, Яндекса, Мегафона и других лидеров рынка, перешедших в пандемию на удаленную работу. Компании открыли массовый доступ к своим учебным курсам, ранее доступным только сотрудникам.

Сервис «**Российская образовательная платформа**» – это сервис, размещенный на образовательной платформе «Университет 2035» (<https://products.2035.university/platform>). Он является инструментом, позволяющим государству стабильно обеспечивать кадрами задачи технологического развития, оперативно реагировать на текущие и будущие потребности, включая новые компетенции в траектории развития граждан на территории всей страны. Более 257 000 человек прошли обучение по сквозным технологиям и цифровым профессиям с 2019 года, 760 000 пользователей заполнили цифровой профиль, чтобы подтвердить свои компетенции для работодателей, более 66 000 единиц образовательного контента размещено на данном сервисе.

Постановлением Правительства РФ от 2 марта 2019 г. № 234 «О системе управления реализацией национальной программы «Цифровая экономика Российской Федерации» и по поручению Президента РФ на базе ведущих университетов страны запланировано открытие более двадцати пяти Центров компетенций в рамках федерального проекта «Оценка и развитие управленческих компетенций в российских образовательных организациях». На базе «Университет 2035» создан **Центр компетенций по кадрам для цифровой экономики** (далее – Центр компетенций), <https://digitalskills.center/>.

Полагаем, что этот опыт актуален и для Беларуси. Центры компетенций следует создавать не только на базе учреждений образования, но и иных организацией, осуществляющих образовательную деятельность.

Согласно паспорту Федерального проекта «Кадры для цифровой экономики» Центр компетенций в 2021-2024 году реализует следующие мероприятия:

- программы дополнительного образования для различных групп населения с целью развития компетенций цифровой экономики;
- образовательные программы по подготовке управленцев и команд цифровой экономики;
- осуществляет независимую оценку компетенций цифровой экономики;
- сервисы поддержки цифровой грамотности и дистанционной занятости;
- обеспечивает координационные, экспертно-аналитические и организационно-технические работы по реализации и мониторингу мероприятий федерального проекта, а также решает иные задачи.

Кроме того, в РФ созданы базовая модель компетенций и сервисы работы с данными по кадрам для цифровой экономики. Например, сервис «**КЛИК-CDO. Подготовка лидеров и команд CDO-менеджеров**» (англ. – *Chief Data Officer*, русс. – руководитель по работе с данными, *CDO*) – программа повышения квалификации в рамках федерального проекта «Кадры для цифровой экономики» <https://clickcdo.ru/>. Формат программы

позволяет осуществлять как обучение цифровым компетенциям и проектной деятельности, так вести и проработку собственных проектов. Целевой аудиторией программы являются – представители федеральной и региональной власти, отвечающие за реализацию национальной программы «Цифровая экономика РФ», руководители и менеджеры крупных федеральных компаний, заинтересованных в создании цифровых решений, отраслевые и научные организации, работающие в сфере развития цифровой экономики. Прежде чем приступить к обучению, предлагается пройти опрос по оценке уровня цифровизации региона, представителем которого является обучающийся. Учеба проходит в дистанционном формате. По итогам выдается документ о дополнительном профессиональном образовании. Более 16 500 человек обучились на программах CDO с 2019 по 2021 гг., более 1 600 проектов по управлению на основе данных разработали команды, проходящие обучение, 85 регионов приняли участие в программе CDO. Проекты разрабатываются командами цифровой экономики федеральных и региональных органов исполнительной власти РФ, госкорпораций, бюджетных и образовательных организаций, промышленных предприятий и коммерческих компаний в рамках акселерационного модуля образовательных CDO-программ. Проекты создаются в процессе командной работы и прорабатываются совместно с наставником. На финальной защите команды представляют свои проекты, эксперты дают оценку проектам. Проекты размещаются в витрине CDO-проектов [6].

Интересен также следующий сервис – «Корпоративная генетика» (*CORPGEN*), облачная система оценки корпоративной культуры для целей организационной трансформации. На основе опроса сотрудников и анализа организационной культуры по 44 параметрам *CORPGEN* может идентифицировать негативные явления и риски в команде.

Следует отметить, что 12 апреля 2021 г. создан и начал работать еще один цифровой университет – Цифровой университет для сообщества ИНТОСАИ (*U-INTOSAI*) – один из ключевых проектов председательства Счетной палаты России в Международной организации высших органов аудита (ИНТОСАИ). Сегодня на сайте *U-INTOSAI* размещены 150 курсов на 5 языках, более 160 уроков, по представленным ниже направлениям (рисунок 2). Слушателями цифрового университета являются около 9 000 человек. Материалы, размещенные на образовательных платформах (предоставляется бесплатно после авторизации), могут быть полезны для самой широкой аудитории: студентов, государственных служащих и всех граждан. Материалы группируются по тематическим направлениям, что облегчает поиск необходимого для изучения курса.



Рисунок 2. Ключевые тематики курсов [7]

Заслуживает внимания опыт работы «**Национальной платформы открытого образования**» (далее – Платформа открытого образования), которая предлагает онлайн-курсы по базовым дисциплинам, изучаемым в российских университетах (<https://openedu.ru>). Платформа открытого образования создана Ассоциацией «Национальная платформа открытого образования», учрежденной ведущими университетами – МГУ имени М.В. Ломоносова, СПбПУ, СПбГУ, НИТУ «МИСиС», НИУ ВШЭ, МФТИ, УрФУ и Университет ИТМО. Все курсы, размещенные на Платформе открытого образования, доступны без формальных требований к базовому уровню образования и доступны для онлайн-обучения. Пройти обучение могут как школьники и студенты, так и работники государственных органов и организаций. До прохождения курса происходит оценка знаний слушателя. По результатам обучения (при желании за оплату) выдается сертификат (при условии прохождения контрольных мероприятий онлайн-курса с идентификацией личности обучающегося и контролем условий их прохождения). Курсы сгруппированы по наименованию высшего учебного заведения, предоставившего курс, по тематическим направлениям, по дате проведения, что удобно для организации поиска информации. Всего – более 1200 курсов, 205 программ повышения квалификации (платно). Форма обучения заочная (дистанционная).

Нельзя не отметить и систему «Цифровой гражданин» – это профессиональная платформа для тестирования и повышения уровня цифровой грамотности для каждого вне зависимости от его уровня знаний (<https://it-gramota.ru/>). На основе диагностики цифровых компетенций строится индивидуальная траектория обучения. Цель – подготовка человека к жизни и работе в цифровой экономике и повышение его конкурентоспособности на рынке труда. Тест состоит из 64 вопросов. Первый шаг – регистрация, после которой тестируемый попадет в свой личный кабинет. В нем будут отражаться результаты тестирования (этой и всех последующих попыток) и рекомендации по дальнейшему обучению, доступный для скачивания сертификат и ссылки на полезные образовательные ресурсы.

Услуги тестирования цифровой грамотности открыты для представителей государственных структур – тест поможет руководителям исполнительных органов власти в разных регионах страны оценивать готовность региона к цифровой трансформации, отслеживать достижение показателей цифровых компетенций населения региона в целом, а также отдельных целевых подгрупп с целью выстраивания индивидуальной, наиболее эффективной траектории обучения. В частности, для компаний и корпораций – тест поможет руководителям и HR-службам оценить готовность сотрудников к цифровой трансформации и сформировать эффективную внутрикорпоративную образовательную стратегию.

Заключение. Таким образом, система профессионального развития кадров для цифровой экономики должна отвечать новым требованиям – обеспечивать эффективное функционирование сервисов электронного правительства, краудсорсинговых платформ, официальных ведомственных веб-ресурсов, групп и страниц в социальных сетях для поддержания диалога и взаимодействия с работниками и населением. Новый функционал работников организаций, вытекающий из процесса диджитализации государственного управления, требует выработки новых подходов к образовательной модели их подготовки, переподготовки, повышения квалификации. На первое место в этой связи выходят навыки владения цифровыми компетенциями, а также постоянное совершенствование *soft skills*, причем компетентностный подход является методологической основой реализации образовательных программ, и соответственно - профессионального развития в условиях цифровизации экономики.

Из проанализированных подходов организации образовательной деятельности в Российской Федерации определено, что наиболее востребованными являются курсы

повышения квалификации, образовательные программы переподготовки, семинары, тренинги, видеоуроки, программы для самостоятельного изучения. Развивать цифровые компетенции работников позволяют онлайн возможности современного образования. В современных условиях повышенных скоростей на смену offline обучению приходит online обучение. Многие образовательные площадки, цифровые университеты внедряют в практику электронное обучение. Дистанционные технологии становятся очень востребованными.

Список литературы

- [1] Паспорт национальной программы "Цифровая экономика Российской Федерации" (утв. президиумом Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам 24 декабря 2018 г. № 16) [Электронный ресурс]. – Режим доступа: // <https://rulaws.ru/acts/Pasportnatsionalnoy-programmy/> – Дата доступа: 20.01.2024.
- [2] Президентская академия запустит онлайн-университет для госслужащих. [Электронный ресурс]. – Режим доступа: <https://www.ranepa.ru/news/prezidentskaya-akademiya-zapustit-onlayn-universitet-dlya-gossluzhashchikh/> . – Дата доступа: 29.01.2024
- [3] Минкомсвязи и Минтруд создадут ресурс для развития госслужащих [Электронный ресурс] . – Режим доступа: <https://www.comnews.ru/content/121623/2019-08-27/minkomsvyazi-i-mintrud-sozdadut-resurs-dlya-razvitiya-gossluzhashchih> . – Дата доступа: 28.01.2024.
- [4] Шуваев, А.В. Траектории развития цифровой экономики и дигитализация образования [Электронный ресурс] // А.В.Шуваев [и др.] . – Режим доступа: <https://cyberleninka.ru/article/n/traektorii-razvitiya-tsifrovooy-ekonomiki-i-digitalizatsiya-obrazovaniya> . – Дата доступа: 24.01.2024.
- [5] Университет 20.35 [Электронный ресурс]. – Режим доступа: <https://www.2035.university/>. – Дата доступа: 28.01.2024.
- [6] Витрина CDO-проектов за 2020–2021 год [Электронный ресурс] . – Режим доступа: <https://cdo.2035.university/projects> . – Дата доступа: 29.01.2024.
- [7] На платформе Цифрового университета ИНТОСАИ опубликован юбилейный 100-й курс [Электронный ресурс] . – Режим доступа: <https://news.myseldon.com/ru/news/index/272948002> . – Дата доступа: 29.01.2024.

Авторский вклад

Авторы внесли равноценный вклад в написании статьи

E-LEARNING FOR HUMAN RESOURCE DEVELOPMENT: THE EXPERIENCE OF THE RUSSIAN FEDERATION

I. Sidorchuk

Deputy Director for Scientific and Methodological Work of the Institute of Information Technologies of BSUIR, Candidate of Legal Sciences, Assistant Professor

A. Okhrimenko

Director of the Institute of Information Technologies of BSUIR, Candidate of Technical Sciences, Assistant Professor

E.Krys

Head of the sector of scientific and methodological work of the Educational and Methodological Department of the Institute of Information Technologies of BSUIR, Master of Management and Law

Abstract. The article analyzes the experience of the Russian Federation in organizing training for managers and specialists to work in the digital economy. Attention is focused on the online opportunities of modern education. It is concluded that many educational platforms and digital universities are putting e-learning into practice. The demand for remote technologies is growing.

Keywords: digital transformation, digital economy, development of digital competencies, personnel for the digital economy, additional professional development.

УДК 004.021:004.75

АЛГОРИТМ ИДЕНТИФИКАЦИИ ТРАНСПОРТНЫХ СРЕДСТВ В УСЛОВИЯХ НЕПОЛНЫХ ДАННЫХ



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, позволяющий подготовить необходимые данные для идентификации транспортного средства, номер которого частично скрыт. Очевидно, что полученные данные доопределяются поисковыми средствами специализированных баз данных, однако представленная работа позволяет организовать два важных подготовительных этапа: распознавание данных по снимкам фотокамер и интерпретацию полученных данных для дальнейшего использования соответствующими регулирующими структурами.

Ключевые слова: транспортная сеть, пропускная способность, распознавание снимка, доопределение идентификатора, поддержка гипотез.

Введение. Непрерывное развитие транспортной инфраструктуры обусловлено взрывным ростом количества средств передвижения и развитием коммуникаций между различными географическими субъектами. Ускоряющиеся темпы роста количества личного транспорта, технических характеристик по скорости передвижения и объективная необходимость передвигаться обуславливают необходимость развития скоростных автомагистралей, что позволяет не только увеличить скорость передвижения и безопасность, но также позволяет осуществлять дальнейшее развития за счет монетизации

дорожных сервисов и развития коммерческих автобанов. Однако, наряду с положительными аспектами внедрения скоростных трасс, существует большая вероятность уклонения водителей от оплаты путём сокрытия части регистрационного номера транспортного средства. Аналогичная проблема возникает даже при наличии средств фотофиксации нарушения правил дорожного движения. Зачастую водитель с частично скрытыми идентификационными данными остается безнаказанным даже при совершении серьезного правонарушения.

С практической точки зрения данная работа не имеет аналогов. На рынке специализированных программных продуктов есть ряд приложений («*Smart PlateReader*», «*Recognizer*»), которые позволяют определить номер автомобиля по данным фотофиксации, однако работать с входным потоком нечетких данных эти системы не способны.

Актуальность. Как правило, нечеткость данных регистрационного номера обусловлена погодными условиями или намеренным сокрытием номера автолюбителем; и в том и другом случае, в объектив фотокамер попадает только часть номера. Очевидно, что с юридической точки зрения, данная информация не может являться основанием для применения мер противодействия совершенным правонарушениям. Данная работа, очевидно, имеет не только научную, но, прежде всего, социальную актуальность. Повышение уровня ответственности водителей является основой безопасной и комфортной дорожной обстановки. С другой точки зрения, повышение уровня точности идентификации правонарушений позволяет сократить нагрузку и/или уменьшить штат сотрудников ГИБДД, в конечном итоге, необходимость патрулирования скоростных трасс может существенно сократиться.

Процедуру идентификации можно укрупнено разделить на следующие этапы:

1 Распознавание снимка фотокамеры для классификации буквенных символов, цифр и относительных размеров «пустого» пространства, не поддающегося классификации.

2 Восстановление и доопределение полученных неполных данных

3 Сопоставление полученной информации с данными официальных баз данных

4 Построение вероятностной модели дальнейшего перемещения транспортного средства

Первый этап (подробно описан в [1]) достаточно успешно реализуется путем применения нейронных сетей, широко представленных в открытых библиотеках. Известная часть регистрационного номера распознается с фотографии, загруженной в программу, методами библиотеки *OpenCV*, далее применяется функционал, предусмотренный методами библиотеки *Keras*, обеспечивающей взаимодействие с искусственными нейронными сетями, в частности, со сверточными сетями глубокого обучения. В результате снимок камеры фотофиксации движения преобразуется в понятный программе набор данных (рисунок 1).

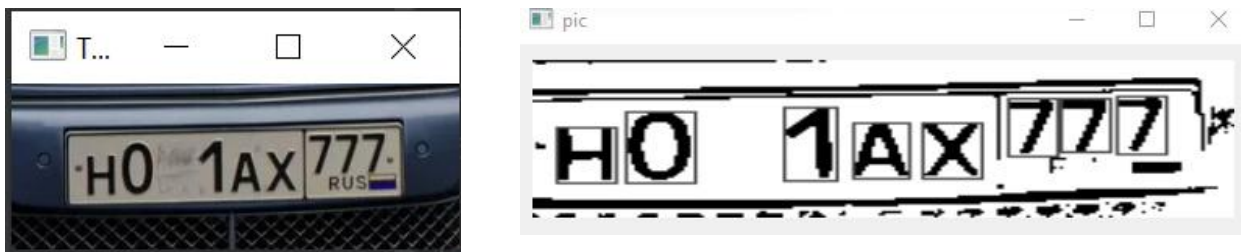


Рисунок 1. Экранные формы входных и выходных данных программного модуля распознавания снимка транспортного средства

Таким образом, первая задача решается достаточно полно.

Алгоритм доопределение полученных неполных данных. В задаче доопределения скрытых цифр номера транспортного средства можно выделить две основные подзадачи:

В качестве исходных данных рассматривается номер с одной скрытой цифрой. Отсутствие одной цифры не вызывает сложностей восстановления. Достаточно хранить сумму цифр номера транспортного средства в БД регулятора и сравнивать с суммой данных, полученных после этапа распознавания снимка (аддитивный критерий поиска числовых данных):

$$j_n = s - (j_1 - j_2), \text{ при ограничениях } \begin{matrix} j_n > 0 \\ 0 < j_n < 10 \end{matrix} = s - (j_1 - j_2)$$

Алгоритмически этот процесс реализуется небольшим программным модулем (рисунок 2).

```
s=[27,23,6,7,10]#суммы цифр номеров "подходящих" машин
i=str(input())# введите известные цифры номера
a=[]
st=0
for j in i:
    st=st+int(j)
for j in s:
    q=j-st
    if q>=0 and len(str(q))==1:
        a.append(q)
print(a)
```

Рисунок 2. Фрагмент программного модуля, реализующего доопределения скрытой цифры номера транспортного средства

В качестве исходных данных рассматривается номер с двумя скрытыми цифрами. Определение пары пропущенных цифр становится более трудоемкой задачей с вычислительной точки зрения. Предложенный выше подход, предусматривающий аддитивный критерий поиска, подразумевает перебор некоторого количества комбинаций закрытых цифр. Алгоритмически этот процесс реализуется программным модулем, фрагмент которого представлен ниже (рисунок 3).

```
s=[27,23,6,7,10]#суммы цифр номеров "подходящих" машин
i=int(input())# введите известную цифру номера
a=[]
r=[]
c1=[0,1,2,3,4,5,6,7,8,9]
for j in s:
    q=j-i
    if 0<=q<=18:
        a.append(q)
for x1 in c1:
    for x2 in c1:
        for j in a:
            if j==x1+x2:
                r.append([x1,x2])
print(r)
```

Рисунок 3. Фрагмент программного модуля, реализующего доопределения двух скрытых цифр номера транспортного средства

В данной работе предлагается ранжировать полученные программные данные по вероятностному критерию с целью увеличения скорости идентификации транспортного средства. До этапа обращения к БД с уточняющей информацией (в представленном

фрагментах программ используется термин «подходящие» машины) по марки и цвету подозрительного транспортного средства предлагается произвести предварительную обработку полученных гипотез о скрытых парах чисел.

Для реализации данного подхода строится частотная матрица M ($M_{i,j} = n_{ij} / N$, где n_{ij} - количество комбинаций i -ой цифры и j -ой буквы относительно общего количества автомобилей в регионе), отражающая все варианты комбинаций букв и цифр в номерах ТС по регистрационным данным ГИБДД соответствующего региона:

$$M = \begin{array}{c|cccccccc} & a & б & в & г & д & е & . & . & . \\ \hline 1 & 0,4 & 0,7 & 0,45 & . & . & . & . & . & . \\ 2 & 0,3 & 0,6 & 0,4 & . & . & . & . & . & . \\ 3 & 0,56 & 0,63 & . & . & . & . & . & . & . \\ 4 & 0,09 & . & . & . & . & . & . & . & . \\ 5 & 0,22 & . & . & . & . & . & . & . & . \\ 6 & 0,25 & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \end{array}$$

Таким образом, для каждой гипотезы можно дополнительно ввести, «поддерживающий» критерий, позволяющий ранжировать полученные ранее результаты распознавания.

Например, для номера $A^{**3}AH$ (* указаны пропущенные цифры) результаты работы алгоритма имели бы следующий вид:

Поддержка	Номер для следующей итерации распознавания
0,56	3*3
0,4	1*3
0,3	2*3
...	...

Предложенный подход для первичного доопределения скрытого номера, подразумевает хранение в БД дополнительной информации, однако, современные технологии позволяют увеличивать объем хранимых данных для обеспечения высокой скорости обработки информации [5].

Заключение. В статье предложены алгоритмы доопределения скрытых цифр номера транспортного средства на основе аддитивного критерия и дополнительной аналитической обработки полученных данных. Вероятностный критерий на основе алгоритма обработки нечетких данных позволяет сократить время поиска информации и уменьшить количество обращений к БД, что значительно уменьшит транзакционную нагрузку на сервер.

В разрезе такого вопроса как безопасность дорожного движения, скорость обработки информации о правонарушении ПДД может повлиять на генерацию ответных мер, вплоть до автоматизации вынесения соответствующих постановлений, что, в конечном счете, позволит повысить ответственность водителей.

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] [3] В.А. Пышный. Моделирование загрузки транспортной сети // Известия ТулГУ. Технические науки. 2012. Вып. 2. С. 457-473.

[3] Гасников А.В., Кленов С.Л., Нурминский Е.А. и др. Введение в математическое моделирование транспортных потоков: учеб. пособие. М.: МФТИ, 2010. — 362 с.

[4] 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

[5] 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.

[6] Корячко В.П., Бакулев А.В., Бакулева М.А. АЛГОРИТМ НАХОЖДЕНИЯ ОПТИМАЛЬНОГО ПУТИ В ТРАНСПОРТНОЙ СЕТИ В УСЛОВИЯХ ИЗМЕНЯЕМЫХ ПАРАМЕТРОВ// В сборнике: BIG DATA И АНАЛИЗ ВЫСОКОГО УРОВНЯ. Сборник научных статей IX Международной научно-практической конференции. В 2 частях. Минск, 2023. С. 161-166.

Авторский вклад

Авторы внесли равноценный вклад в написании статьи

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

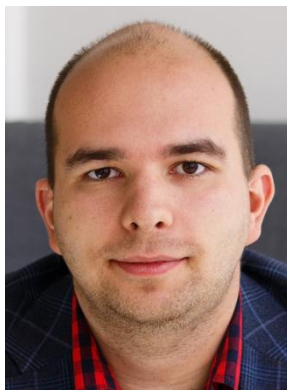
*Associate Professor of the
Department of Computer Aided
Design of Computing Facilities
Ryazan State Radio Engineering
University named after V.F.
Utkin (RSREU),
PhD of Technical Sciences,
Associate Professor*

Abstract. This research devote to develop the approach to identifying number of the car, if some information about one is hidid. The photos are determined by search tools of specialized databases. For this purpose is used deep learning neural network technology and specialized open library like Keras. Two base tasks are solved: number recognition from camera images and interpretation of the obtained data. Fast recognition allows increase quality information about break law and increase safety transport networks.

Keywords: transport network, throughput, image recognition, deep identification, the hypothesis support.

УДК 004.021:004.75

ИЗУЧЕНИЕ АВТОМАТИЗИРОВАННЫХ СИСТЕМ КОНТРОЛЯ И УПРАВЛЕНИЯ ГОРОДСКИМИ ФЕРМАМИ ДЛЯ РАЗРАБОТКИ ТЕХНИЧЕСКОЙ КОНЦЕПЦИИ СИСТЕМЫ УПРАВЛЕНИЯ ГОРОДСКИМИ ФЕРМАМИ В РЕЖИМЕ ОНЛАЙН



А.К. Куликов

*Доцент кафедры приборов и информационно-измерительных систем РТУ МИРЭА,
кандидат технических наук
science.andrey.kulikov@gmail.com*

А.К. Куликов

Окончил Федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет». Область научных интересов связана с разработкой методов и алгоритмов построения информационно-измерительных систем, организацией учебного и научно-исследовательского процессов в техническом университете.

Аннотация. Проведен обзор и анализ современных научных исследований и публикаций на тему создания автоматизированных систем управления городскими фермами. Рассмотрены наиболее распространенные архитектуры таких систем. Разработана техническая концепция системы управления городской фермой с использованием методов машинного обучения и автоматизации процессов (*ML-Ops*) для обеспечения управления в режиме онлайн.

Ключевые слова: *ML-ops*, сити-фермы, системы управления.

Введение. На текущий момент задача обработки данных для генерации точных предсказаний объема выпускаемой продукции с квадратного метра умного фермерского хозяйства является актуальной [1-6]. Особенно ярко об этом может свидетельствовать развитие стартапов: *InFarm*, *Farm One*, *Gotham Greens* и др. Специалисты в области исследования стремятся определить, какие параметры имеют наибольшее влияние на процесс роста различных типов растений, с целью создания наиболее точного прогноза. Для выполнения обработки данных и формирования прогнозных моделей на их основании создаются различные системы управления [7-15]. Основными задачами подобных систем управления, являются: получение и хранение информации, обработка информации, оповещение оператора о результатах обработки, коррекция среды по результатам обработки.

Анализ систем управления смарт-теплицами. По результатам анализа информации из открытых источников на сегодняшний день системы управления представляют из себя Системы разделенные на четыре логических уровня, которые представлены на рисунке 1.

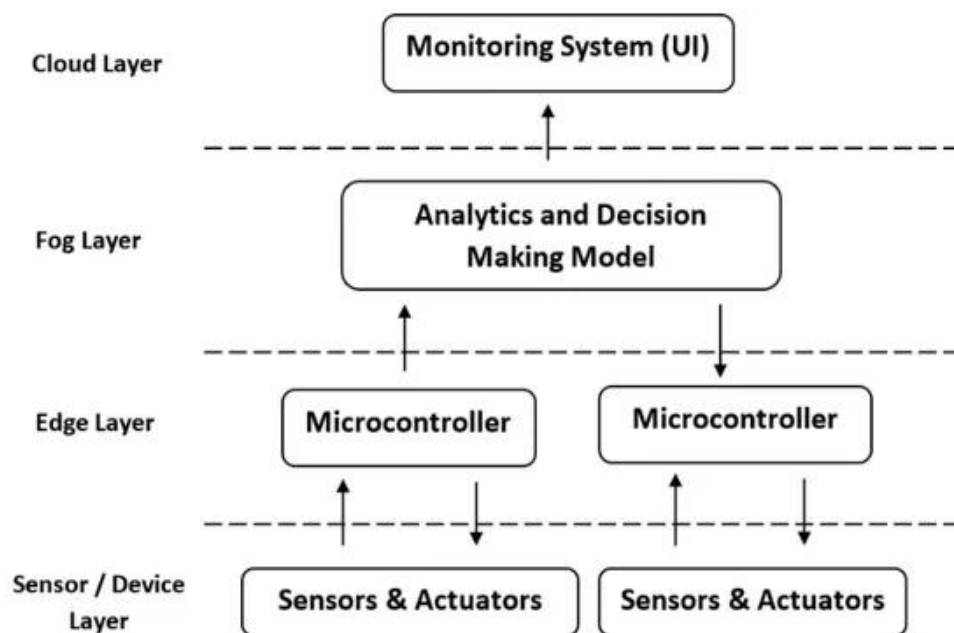


Рисунок 1. Логические уровни системы управления смарт-теплицами.

Таким образом, структура системы управления включает в себя следующие уровни: сенсоры, обработка информации с сенсоров, анализ информации и принятие решений, мониторинг функционирования. Зачастую системы подключаются на уровне сенсоров и на уровне обработки информации при помощи *Wi-Fi*, *LoRaWan*, *NB-IoT* или аналогичных беспроводных и проводных сетей. Сами сенсоры, либо изготавливаются в лабораториях самостоятельно, либо покупаются. Сенсоры собирают данные по параметрам, необходимым для работы прогнозирующих моделей. На этапе обработки данных с сенсоров используются микроконтроллеры с архитектурой *ARM* (*Arduino* или *Raspberry Pi*). На уровне принятия решений происходит планирование, диагностика и прогнозирование. Программная и аппаратная реализации данного уровня выполняются чаще всего самостоятельно научными коллективами, обычно это включает в себя использование протокола связи *MQTT* для передачи данных с предыдущих уровней, а также *SQL*-запросы к базе данных на основе полученных данных. Аппаратная реализация этого уровня чаще всего выполняется на базе *Raspberry Pi* и аналогов на микроконтроллерах с *ARM* архитектурой, ноутбуках или отдельных компьютерах. Уровень мониторинга системы в целом программно реализован в готовых инструментах мониторинга для решений *IoT* или изготовленных самостоятельно.

Наиболее интересными являются исследования, проводимые *Wageningen University & Research*, приблизительный вид их системы управления на рисунке 2. На текущий момент этим научно-исследовательским коллективом организованы состязания по созданию точных прогнозирующих моделей за счет предоставления научным коллективам доступа к своим смарт-фермам, а также возможности анализировать в режиме реального времени полученные результаты и корректировать их. Помимо этого, по результатам их исследований существует портал, который позволяет пользователям на основе своих данных получить прогнозы по выращиванию собственных культур.

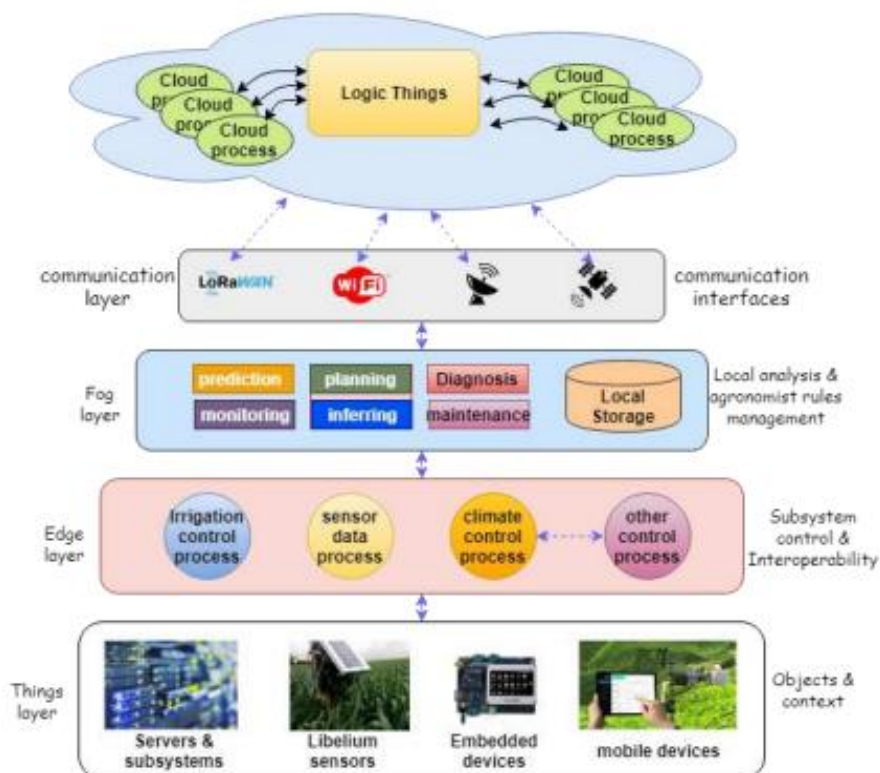


Рисунок 2. Wageningen University & Research система управления.

Кроме того, в контексте автоматизации городских ферм следует отметить проект ученых из Саудовской Аравии на рисунке 3. Они предложили конкретные рекомендации по техническому облику системы контроля. Основной целью этой системы является наблюдение за условиями внутри теплицы и контроль внутренней температуры с целью снижения потребления энергии при сохранении благоприятных условий для повышения урожайности. В качестве инструмента для мониторинга условий в теплице и определения оптимальной эталонной температуры используется модель сетей Петри. Температура затем передается в блок управления температурой. Вторая цель – разработка энергоэффективной масштабируемой системы для обработки больших объемов данных, поступающих от датчиков. Для этого используется динамическая графическая модель данных, которая может быть использована для прогнозирования производства, скорости роста урожая, энергопотребления и других связанных вопросов в будущем. Проект предусматривает унификацию различных форматов неструктурированных данных, полученных от разных типов устройств интернета вещей, с помощью преобразований моделей и моделируемой архитектуры. Это позволяет преобразовывать данные в структурированный вид.

Таким образом, был сделан вывод о важности реализации в техническом облике блока быстрого тестирования *ML* моделей и их быстрого апробирования на реальных растениях с целью получения точной модели, которую можно было бы передавать в другие сити-фермы для масштабирования результата выхода продукции конкретной сельскохозяйственной культуры.

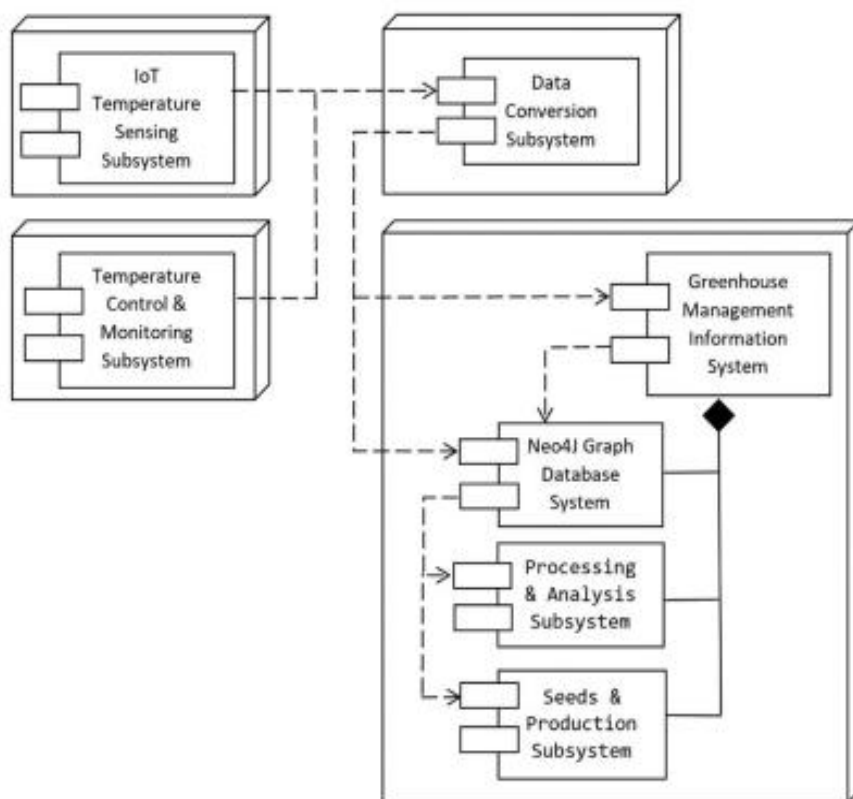


Рисунок 3. Интеллектуальная система на базе информационных технологий для контроля температуры в теплице от ученых из Саудовской Аравии.

Технический облик системы управления сити-фермой. Таким образом, в результате изучения материалов из открытых источников и анализа современных систем контроля, основанных на принципах *DevOps*, было решено разработать систему, в основе которой будет лежать технология *ML-Ops*. Данный шаг, позволит не только выполнять исследования в режиме реального времени для подбора моделей машинного обучения в соответствии с получаемыми данными и культурами, находящимися в городской ферме, но и выполнять интеграцию новых моделей и развертывание старых моделей в различных сценариях.

Аппаратная реализация такой системы подразумевает принятие ряда решений:

- Использование сенсоров и датчиков, работающих по протоколу *I2C* или *1-Wire*;
- Плата для обработки и отправки информации на сервер (*NodeMCU* или аналоги), способная передавать данные по *Bluetooth*, *Wi-Fi* или кабелю;
- Протокол для передачи данных от узлов к серверу (*MQTT*);
- Реализация уровней обработки и мониторинга в форме облачного приложения для быстрого масштабирования с дублированием серверной части.

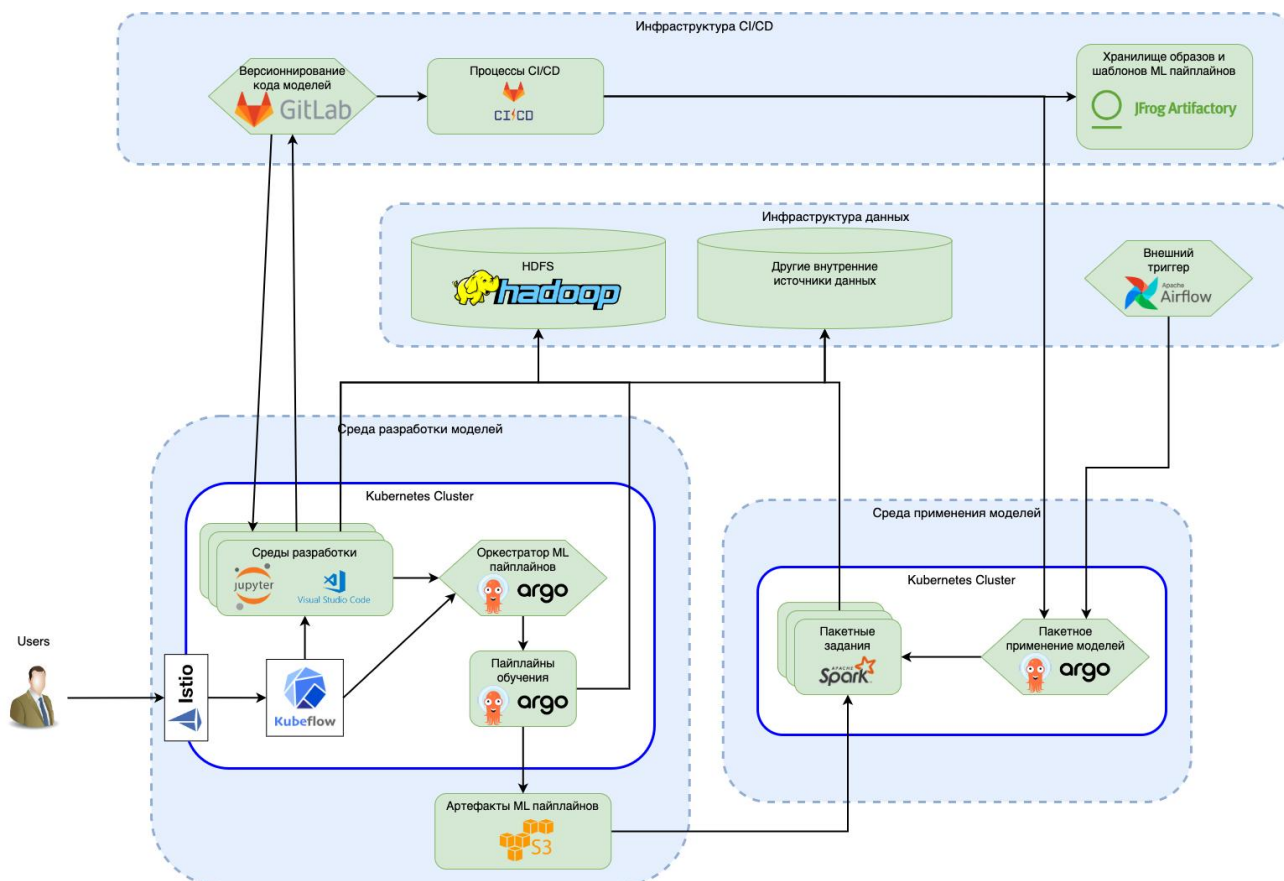


Рисунок 4. Технический облик системы управления смарт-теплицами.

Заключение. В ходе исследования была разработана концепция системы контроля городских ферм, которая позволяет избежать необходимости самостоятельной разработки инфраструктуры системы управления и сконцентрироваться на уровнях сенсоров и обработки данных с сенсоров.

В результате реализации данной концепции, как показано на рис. 4, ожидается получение следующих преимуществ: автоматизация сохранения данных, управление версиями моделей, агрегирование наборов данных, тестирование новых моделей с помощью экспериментов и новых данных, визуализация данных.

Список литературы

- [1] Kim, W.; Soon, B.M. Advancing Agricultural Predictions: A Deep Learning Approach to Estimating Bulb Weight Using Neural Prophet Model. *Agronomy* 2023, 13, 1362. <https://doi.org/10.3390/agronomy13051362>
- [2] Thimmegowda, M.N.; Manjunatha, M.H.; Huggi, L.; Shivaramu, H.S.; Soumya, D.V.; Nagesha, L.; Padmashri, H.S. Weather-Based Statistical and Neural Network Tools for Forecasting Rice Yields in Major Growing Districts of Karnataka. *Agronomy* 2023, 13, 704. <https://doi.org/10.3390/agronomy13030704>.
- [3] Wang, Y.-H.; Su, W.-H. Convolutional Neural Networks in Computer Vision for Grain Crop Phenotyping: A Review. *Agronomy* 2022, 12, 2659. <https://doi.org/10.3390/agronomy12112659>
- [4] Zhou, J.; Wang, B.; Fan, J.; Ma, Y.; Wang, Y.; Zhang, Z. A Systematic Study of Estimating Potato N Concentrations Using UAV-Based Hyper- and Multi-Spectral Imagery. *Agronomy* 2022, 12, 2533. <https://doi.org/10.3390/agronomy12102533>
- [5] Liu, C.; Zhao, C.; Wu, H.; Han, X.; Li, S. ADDLight: An Energy-Saving Adder Neural Network for Cucumber Disease Classification. *Agriculture* 2022, 12, 452. <https://doi.org/10.3390/agriculture12040452>
- [6] Pham, Q.V.; Nguyen, T.T.N.; Vo, T.T.X.; Le, P.H.; Nguyen, X.T.T.; Duong, N.V.; Le, C.T.S. Applying the SIMPLE Crop Model to Assess Soybean (*Glicine max. (L.) Merr.*) Biomass and Yield in Tropical Climate Variation. *Agronomy* 2023, 13, 1180. <https://doi.org/10.3390/agronomy13041180>

- [7] Jin, X.-B.; Zheng, W.-Z.; Kong, J.-L.; Wang, X.-Y.; Zuo, M.; Zhang, Q.-C.; Lin, S. Deep-Learning Temporal Predictor via Bidirectional Self-Attentive Encoder–Decoder Framework for IOT-Based Environmental Sensing in Intelligent Greenhouse. *Agriculture* 2021, 11, 802. <https://doi.org/10.3390/agriculture11080802>
- [8] Zhang C, Liu Z. Application of big data technology in agricultural Internet of Things. *International Journal of Distributed Sensor Networks*. 2019;15(10). doi:10.1177/1550147719881610
- [9] Jirapond Muangprathub, Nathaphon Boonnam, Siriwan Kajornkasirat, Narongsak Lekbangpong, Apirat Wanichsombat, Pichetwut Nillaor, IoT and agriculture data analysis for smart farm, *Computers and Electronics in Agriculture*, Volume 156, 2019, Pages 467-474, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2018.12.011>.
- [10] Haixia Li, Yu Guo, Huajian Zhao, Yang Wang, David Chow, Towards automated greenhouse: A state of the art review on greenhouse monitoring methods and technologies based on internet of things, *Computers and Electronics in Agriculture*, Volume 191, 2021, 106558, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2021.106558>.
- [11] Hemming, S.; de Zwart, F.; Elings, A.; Righini, I.; Petropoulou, A. Remote Control of Greenhouse Vegetable Production with Artificial Intelligence—Greenhouse Climate, Irrigation, and Crop Production. *Sensors* 2019, 19, 1807. <https://doi.org/10.3390/s19081807>
- [12] A. F. Subahi and K. E. Bouazza, "An Intelligent IoT-Based System Design for Controlling and Monitoring Greenhouse Temperature," in *IEEE Access*, vol. 8, pp. 125488-125500, 2020, doi: 10.1109/ACCESS.2020.3007955.
- [13] Hemming, S.; Zwart, F.d.; Elings, A.; Petropoulou, A.; Righini, I. Cherry Tomato Production in Intelligent Greenhouses—Sensors and AI for Control of Climate, Irrigation, Crop Yield, and Quality. *Sensors* 2020, 20, 6430. <https://doi.org/10.3390/s20226430>
- [14] Maraveas, C. Incorporating Artificial Intelligence Technology in Smart Greenhouses: Current State of the Art. *Appl. Sci.* 2023, 13, 14. <https://doi.org/10.3390/app13010014>
- [15] Rokade, A.; Singh, M.; Malik, P.K.; Singh, R.; Alsuwian, T. Intelligent Data Analytics Framework for Precision Farming Using IoT and Regressor Machine Learning Algorithms. *Appl. Sci.* 2022, 12, 9992. <https://doi.org/10.3390/app12199992>

Авторский вклад

Куликов Андрей Кириллович – полная подготовка доклада

QUALITY EVALUATION OF INFORMATION TRANSFER IN A DISPATCHING SYSTEM BASED ON MQTT ARCHITECTURE

A.K. Kulikov

*Associate Professor of the
Department of Devices and
Information and Measuring
Systems of RTU MIREA, PhD of
Technical Sciences*

Abstract. A review and analysis of modern scientific research and publications on the creation of automated management systems for urban farms has been carried out. The most common architectures of such systems are considered. A technical concept of an urban farm management system using machine learning and process automation (ML-Ops) methods to provide online management has been developed.

Keywords: ML-Ops, city forms, information and measurement control systems.

УДК 1256:12

BIG DATA И ADVANCE ANALYTICS В ПЕРСОНАЛИЗИРОВАННОЙ МЕДИЦИНЕ



Н.А. Ларченко
Студент факультета
компьютерного
проектирования кафедры
электронной техники и
технологий БГУИР
nikitadeve@gmail.com



И.В. Андриалович
старший преподаватель
кафедры инженерной
психологии и эргономики,
магистр техники и технологии
andryinna@bsuir.by

Н.А. Ларченко

Окончил Лицей №1 г. Минска при БГУИР. Область научных интересов связана с использованием BIG DATA и Advance Analytics в медицинской сфере.

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.

Аннотация. В статье рассматривается применение технологий *BIG DATA* и анализа больших данных (*Advance Analytics*) для персонализации медицины. Приводятся примеры использования *BIG DATA* в различных областях: от разработки персонализированных лекарств до создания индивидуальных программ по питанию и тренировкам. Обсуждаются перспективы применения *BIG DATA* для мониторинга здоровья пациентов, прогнозирования рисков развития заболеваний, персонализированной терапии в онкологии. Делается вывод о том, что анализ больших данных - ключ к медицине будущего.

Ключевые слова: *BIG DATA*, *Advance Analytics*, персонализированная медицина, персонализированные лекарства, персонализированная диетотерапия, индивидуальные фитнес-программы, персонализированная терапия рака, мониторинг здоровья, прогнозирование рисков, индивидуальный подход.

Введение. В условиях цифровой трансформации общества использование ИКТ в профессиональной деятельности медицинских работников чрезвычайно важно. Информатизация здравоохранения в Республике Беларусь включает в себя внедрение соответствующих программных комплексов в систему здравоохранения как на государственном уровне, так и в отдельных лечебно-профилактических учреждениях.

Технологии больших данных и анализ больших данных (*Advance Analytics*) становятся важными инструментами для выполнения поставленных задач. Они позволяют учитывать особенности каждого пациента при разработке и назначении лечения. Их применение позволяет обнаружить закономерности и создать индивидуальную терапию на основе анализа огромного количества разнообразных данных о пациентах.

В этой статье мы поговорим более подробно о том, как *Big Data* и *Advance Analytics* уже используются для создания индивидуальных подходов к лечению.

BIG DATA и Advance Analytics. Когда речь идет о стремительно растущем объеме разнообразных данных, которые генерируются в современном цифровом мире, термин «*BIG DATA*» используется. Это данные из датчиков, мобильных устройств, социальных сетей и камер видеонаблюдения. Объем, разнообразие форматов и высокая скорость обновления являются основными характеристиками больших данных.

Напротив, анализ больших данных (*Advance Analytics*) – это методы анализа больших данных, которые позволяют извлечь скрытые знания из огромных объемов данных. В первую очередь это технологии машинного обучения и искусственного интеллекта. [1] С их помощью можно построить прогнозные модели, найти сложные закономерности и корреляции и получить новые идеи для принятия решений, основанные на анализе больших данных.

Рассмотрим, как персонализировать лечение с помощью больших данных и передовых аналитических методов.

Персонализированные препараты. Для разработки новых лекарственных препаратов, наиболее подходящих каждому пациенту, наиболее перспективным направлением является использование больших данных. Для определения наиболее безопасного и эффективного лекарства для каждого человека можно использовать данные о геномах, протеомах, метаболомах и других аспектах метаболизма человека.

Например, фармацевтическая компания *23andMe* (США) использует *технологии BIG DATA* для разработки новых противораковых препаратов, которые учитывают генетические характеристики пациентов. Это позволяет значительно повысить эффективность лечения для каждого отдельного пациента. Компания *Biogen* (США), которая разрабатывает новые препараты для лечения нейродегенеративных заболеваний на основе анализа генетических данных, является дополнительным примером.

Китай является мировым лидером в области персонализированной медицины. Компания *BGI* из Китая была пионером в области использования геномных технологий в здравоохранении. В настоящее время *BGI* работает с медицинскими учреждениями по всему миру, чтобы разрабатывать индивидуальные методы диагностики и лечения с использованием *BIG DATA* [2].

Федеральный научно-клинический центр детской гематологии, онкологии и иммунологии имени Дмитрия Рогачева является одним из ведущих в России в области персонализированной медицины. В этом регионе методы таргетной терапии и иммунотерапии, основанные на геномных исследованиях, активно используются для лечения онкологических больных [3]. Республиканский научно-практический центр онкологии и медицинской радиологии в Беларуси занимается разработкой персонализированной медицины. Ученые используют генетическое профилирование опухолей пациентов для подбора таргетной терапии.

Помимо фармкомпаний, технологические гиганты активно участвуют в разработке персонализированных лекарств. Таким образом, *Google* сотрудничает с фармацевтическими компаниями в проекте по использованию машинного обучения для анализа медицинских данных, чтобы создать новые противораковые препараты, подходящие для каждого пациента.

Применение технологий редактирования генома, таких как *CRISPR/Cas9*, привело к значительным прогрессам в этой области. Ученые используют профиль геномов каждого пациента для разработки генно-терапевтических лекарств. Это позволит лечить множество генетических заболеваний.

Прогнозирование рисков. С помощью анализа больших данных можно более точно прогнозировать вероятность развития различных заболеваний у конкретного пациента.

Клиника Мэйо разработала алгоритм на основе искусственного интеллекта, который может прогнозировать риск внезапной смерти с помощью электрокардиограммы. Стартап *Human Longevity* оценивает риск онкологических и сердечно-сосудистых заболеваний на основе данных о образе жизни и генетике.

Выявление генетических предрасположенностей к различным заболеваниям является важным направлением. Уже сейчас существуют тесты, которые оценивают вероятность развития рака молочной железы, яичников и толстой кишки на основе анализа генов *BRCA1* и *BRCA2*, среди прочего. В будущем будут разработаны тесты для различных видов онкологии.

В Беларуси с мая 2023 года внедрён в медицинские центры анализ крови «*PanTum Detect*», который позволяет выявить больше 60 видов рака, что важно, даже на нулевых стадиях. Результаты исследования крови обрабатываются с помощью программы искусственного интеллекта. При положительном результате анализа пациент будет безотлагательно направлен в рамках программы «*PanTum Detect*» на прием в государственную онкологическую клинику к врачу-онкологу. Консультация по результатам анализа и диагностика проводится бесплатно, а лечение проходит на базе РНПЦ онкологии и медицинской радиологии им. Н. Н. Александрова и в областных онкодиспансерах.

Прогнозирование ответа пациента на определенное лечение, в том числе лекарственные препараты, является еще одной важной задачей. Уже разработаны тесты на основе фармакогенетики, которые могут прогнозировать безопасность и эффективность многих популярных лекарств в зависимости от генетических характеристик человека.

Прогностические модели, основанные на сотнях факторов, от генома до образа жизни, могут быть созданы с помощью больших данных и методов машинного обучения. Это позволит более точную оценку риска развития различных заболеваний у конкретного человека. В соответствии с этим будут создаваться индивидуальные программы профилактики и ранней диагностики.

Персонализированная терапия рака. Онкология является одной из наиболее перспективных областей использования больших данных. Сбор и анализ данных об опухоли и организме пациента помогают выбрать оптимальное лечение для каждого отдельного пациента.

Онкологический центр Мемориального онкологического госпиталя Слоуна-Кеттеринга активно использует геномное профилирование для выбора наиболее персонализированной терапии для больных раком. Это обеспечивает более высокие показатели выживаемости.

Foundation Medicine, компания, проводящая тестирование опухолей с использованием технологий *NGS* для подбора таргетных препаратов, представляет собой дополнительный пример. По сравнению со стандартной химиотерапией это значительно повышает эффективность лечения.

Республиканский научно-практический центр онкологии и медицинской радиологии им. Н.Н. Александрова играет важную роль в разработке персонализированных подходов в онкологии в Беларуси.

В центре создан банк опухолевых образцов, что позволяет проводить молекулярно-генетические тесты для подбора таргетных препаратов. Кроме того, с помощью специальных аналитических чипов есть возможность оценить чувствительность опухолей к различным видам химиотерапии. Это увеличивает эффективность лечения онкобольных пациентов.

На базе Центра ведутся исследования, направленные на создание новых биомаркеров, которые могут быть использованы для ранней диагностики и

прогнозирования течения рака. Кроме того, изучаются генетические предикторы эффективности иммунотерапии, что важно для адаптации терапии.

Перспективным направлением лечения рака является применение лекарств на основе мРНК. Эти лекарства стимулируют иммунную систему пациента для борьбы с опухолью. Данные о геноме и протеоме пациента могут помочь им адаптироваться.

Мониторинг состояния здоровья. Современные устройства и приложения, которые отслеживают параметры, такие как пульс, сон, активность и другие, предлагают широкий спектр возможностей для индивидуальной медицинской помощи. Врачи могут более точно оценивать состояние пациента и назначать наилучшую терапию, собирая и анализируя эти данные в динамике.

Например, *Apple* активно интегрирует функции мониторинга здоровья в свои телефоны и часы. Пользователь может отслеживать свое качество сна и частоту пульса, уровень физической активности. Врач может автоматически получать эти данные для мониторинга состояния пациента и корректировки лечения.

Cardio – приложение, которое измеряет артериальное давление и частоту пульса с помощью камеры смартфона. Это позволяет пациентам с сердечно-сосудистыми заболеваниями постоянно следить за важными показателями.

Многочисленные устройства для мониторинга состояния здоровья человека активно разрабатываются и внедряются, помимо смартфонов и умных часов. К примеру, для контроля работы сердца и легких используются специальные датчики, которые крепятся на груди пациента.

Импланты и «умные» пластыри, которые могут собирать информацию о состоянии внутренних органов пациента и передавать ее лечащему врачу с помощью *Bluetooth* или других беспроводных технологий, представляют собой перспективное направление. Это позволяет врачам постоянно отслеживать важные показатели и своевременно реагировать, если они ухудшаются [4].

Устройства виртуальной и дополненной реальности также широко используются для мониторинга состояния пациентов. Например, *MindMaze* создал *VR*-шлем, который отслеживает движения глаз, мимику и позу пациента и передает эти данные врачу для реабилитации пациентов после инсульта. Кроме того, *Doppel* считывает биометрические данные на запястье и помогает снизить тревожность и стресс.

В будущем носимые устройства, датчики и импланты смогут отслеживать множество показателей организма пациента в режиме реального времени, что откроет новые возможности для проактивной и превентивной персонализированной медицины.

Персонализированная диетотерапия. Разработка индивидуальных рекомендаций по питанию для профилактики и лечения различных заболеваний является еще одним важным направлением использования *BIG DATA*.

Например, стартап *DayTwo* может, на основе исследования микробиома кишечника человека, давать подробные советы о том, какие продукты и в каком количестве будут полезны конкретному пациенту. Это позволяет контролировать такие заболевания, как диабет, ожирение и синдром раздраженного кишечника, эффективно.

Другой пример – стартап *Nutrino*, который использует ИИ для анализа данных о генетике, активности, рационе и состоянии кишечника пациента. На основе этих данных разрабатывается индивидуальная диета и рекомендации по питанию, которые помогут улучшить пищеварение, уменьшить вес и улучшить общее состояние здоровья.

Такие методы также широко используются в диетотерапии онкологических больных. Генетические особенности пациента и данные о составе микрофлоры определяют подходящее питание, чтобы уменьшить побочные эффекты и повысить эффективность лечения.

Индивидуальные фитнес-программы. В настоящее время все больше компаний используют большие данные для разработки индивидуальных программ фитнеса и тренировок.

Программа упражнений *Fitbod* использует данные о предыдущих тренировках пользователя, его физическую форму и предпочтения. Приложение *Lark* оценивает образ жизни и предлагает идеи для активности [5].

После травм и операций персонализированный подход также важен. Приложения используют данные о состоянии человека для разработки персонализированных программ восстановления.

Многие тренажеры и фитнес-браслеты записывают свои тренировки. Это помогает в анализе и оптимизации индивидуальных программ.

Использование виртуальной и дополненной реальности является перспективным направлением. Приложения *VR* позволяют создавать индивидуальные тренировки с нагрузкой, соответствующей потребностям клиента. Например, сенсоры отслеживают дыхание и пульс пользователя в реальном времени.

Некоторые компании производят умные тренажеры, которые анализируют состояние здоровья человека и адаптируют свои тренировочные программы. Это предотвращает перенапряжение.

В будущем фитнес будет полностью индивидуализирован на основе обширного профиля пользователя, который включает результаты медицинских и генетических тестов. Это сделает тренировки безопасными и эффективными.

Заключение. Таким образом, более персонализированная медицина может быть достигнута благодаря технологиям *Big Data* и *Advance Analytics*. Их использование на всех этапах, от разработки лекарств до мониторинга состояния пациента, позволяет собирать всю информацию о здоровье пациента, его образе жизни и привычках, строить гипотезы, предлагать алгоритмы обследования, лечения, подбирать врачей и клиники, создавать индивидуальные подходы к профилактике, диагностике и лечению заболеваний, а также улучшать и ускорить работу врачей. Одновременно появляется возможность предупреждать заболевания, что очень важно для дорогостоящих систем здравоохранения. Медицинское будущее зависит от эффективного использования больших данных.

Список литературы

- [1] <https://online.norwich.edu/academic-programs/resources/what-is-advanced-analytics-example-uses-and-applications>
- [2] НМИЦ ДГОИ им. Дмитрия Рогачёва. – [Электронный ресурс] – URL: <https://fnkc.ru/index.jsp?load=aboutcenter> (дата обращения 10.02.2024)
- [3] РНПЦ онкологии и медицинской радиологии им. Н. Н. Александрова. – [Электронный ресурс] – URL: <http://www.oncology.by/development.html>
- [4] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7016513/pdf/sensors-20-00343.pdf>
- [5] <https://lark.com/blog/how-lark-personalizes-for-you>

Авторский вклад

Андрялович Инна Владимировна – руководство исследованием BIG DATA и Advance analytics в персонализированной медицине, постановка задачи исследования.

Ларченко Никита Александрович – описание принципа работы BIG DATA и Advance analytics, формирование структуры статьи, анализ применения технологий BIG DATA и анализа больших данных (Advance Analytics) для персонализации медицины в Республике Беларусь и других странах.

BIG DATA AND ADVANCE ANALYTICS IN PERSONALIZED MEDICINE

N.A. Larchenko

*Student of the Computer Design
Faculty of the Department of
Electronic Engineering and
Technologies BSUIR*

I.V. Andrialovich

*Senior Lecturer of the Department
of Engineering Psychology and
Ergonomics,
master of engineering and
technology*

Annotation. The article discusses the application of BIG DATA and big data analytics (Advance Analytics) technologies for personalizing medicine. Examples of using BIG DATA in various fields are given: from the development of personalized medicines to the creation of individual nutrition and exercise programs. The prospects of BIG DATA application for patient health monitoring, disease risk prediction, and personalized therapy in oncology are discussed. It is concluded that big data analysis is the key to medicine of the future.

Keywords: BIG DATA, Advance Analytics, personalized medicine, personalized medicines, personalized nutrition therapy, individual fitness programs, personalized cancer therapy, health monitoring, risk prediction, individual approach.

УДК 612.76:612.2

АЛГОРИТМ ОБРАБОТКИ ДАННЫХ БИОМЕХАНИКИ ДЫХАТЕЛЬНЫХ ДВИЖЕНИЙ



И.И. Ревинская
ассистент кафедры
электронной техники и
технологии БГУИР
irewinskaya.bsuir@gmail.com



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



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

И.И. Ревинская

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

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.

Е.И. Лещевич

Окончил Белорусский государственный университет информатики и радиоэлектроники. Аспирант кафедры электронной техники и технологии. Область научного интереса – влияние инфразвука на биологические ткани.

Аннотация. Разработан аппаратно-программный комплекс для исследования биомеханических дыхательных движений стенок грудной клетки и живота, а также алгоритм исследования и обработки данных биомеханики дыхания. Предложена методика обработки данных исследований.

Ключевые слова: дыхание, биомеханика, видеокамера, спирометрия, объем легких, алгоритм обработки данных

Введение. Функциональные методы исследования приобретают ключевое значение в диагностике заболеваний легких. Это обусловлено потребностью клиницистов в объективной количественной и качественной характеристике возникающих у больного нарушений дыхания. Своевременная диагностика, наблюдение и лечение пациента может гарантировать эффективность лечения и контроль болезни [1-2]. Пациенты с хроническими респираторными заболеваниями должны следить за состоянием своего дыхания, и любые внезапные изменения в состоянии должны быть немедленно устранены для предотвращения дальнейших обострений.

Классические методы исследования функции внешнего дыхания (спирометрия, пневмотахометрия, бодиплетизмография и т.д.) позволяют выявлять возможные патологии функции дыхания, респираторные заболевания, такие как хроническая обструктивная болезнь легких (ХОБЛ), астма, бронхит, нарушение бронхиальной проводимости и т.д. [2-3]. Спирометрия предназначена для измерения объема вдыхаемого и выдыхаемого воздуха и считается «золотым стандартом» диагностики хронической обструктивной болезни лёгких [3].

В настоящее время все более актуально использовать как контактные, так и бесконтактные методы исследования функции внешнего дыхания. Последние научные исследования в данной области обосновывают возможность применения методов, основанных на анализе движения грудной клетки и живота [3-4]. Различают три подхода к анализу дыхания: на основе записи деформации грудной клетки, вызванной дыхательной активностью (датчики деформации); на основе изменения трансторакального импеданса (датчики импеданса); на основе записи движений грудной клетки (например, наклон грудной клетки, ускорение и скорости) с помощью акселерометра, гироскопа, магнитометра.

Эти методы позволяют косвенно оценивать объемные параметры дыхания по полученным значениям с датчиков (изменение импеданса грудной клетки во время дыхания, изменение емкости датчика, сопротивления и т.д.) из-за движения грудной клетки и брюшной полости. Таким образом, актуально и перспективно разрабатывать методы исследования функции внешнего дыхания по движению грудной клетки и живота. Такие методы позволят решить описанные ранее недостатки.

Целью данной работы является разработка алгоритма обработки данных биомеханики дыхания, позволяющего по значениям отклонений стенок грудной клетки и живота оценивать легочные объемы дыхания.

Материалы и методы. Для исследования биомеханических движений стенок грудной клетки и живота во время дыхания предложен аппаратно-программный комплекс (рисунок 1), который включает в себя видеокамеру, спирометр и персональный компьютер [5-6].

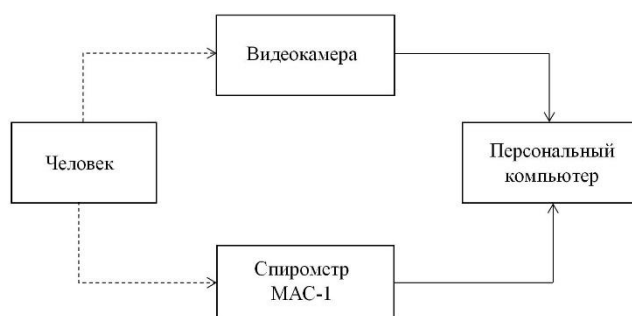


Рисунок 1. Структурная схема аппаратно-программного комплекса

Видеокамера необходима для записи движений стенок грудной клетки и живота испытуемого (субъект) во время дыхания. Спирометр является эталонным прибором измерения объемов вдыхаемого и выдыхаемого воздуха и предназначен для оценки достоверности предложенного способа измерения.

Методика проведения эксперимента приведена на рисунке 2.

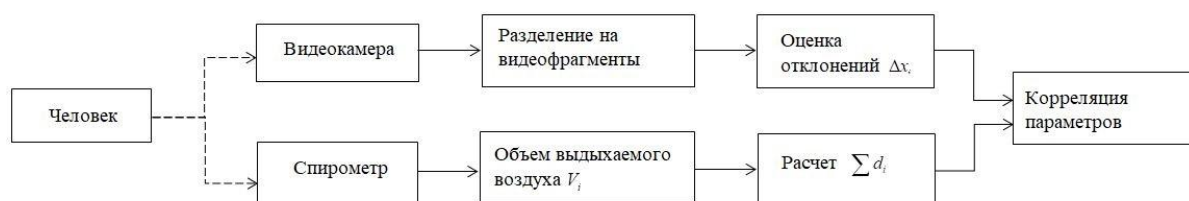


Рисунок 2. Методика проведения эксперимента

На персональном компьютере (ПК) с помощью специализированного программного обеспечения проводится разделение видеофайлов на видеофрагменты, обработка видеофрагментов, извлечение данных со спирометра, а также статистический анализ полученных параметров экспериментальных данных.

Алгоритм обработки данных. Алгоритм регистрации биомеханики дыхания (рисунок 3) основан на способе регистрации биомеханики грудной клетки и живота [5] и аппаратно-программного комплекса [6], который включает в себя следующие этапы [7]:

1 Установка видеокамеры и калибровочной доски.

Калибровочная доска имеет цену деления 0,5 и 1 см по оси абсцисс и 2 см по оси ординат и предназначена для измерения координат крайних точек стенок спины и грудной клетки (живота). Калибровочная доска крепится на стене слева от испытуемого. Видеокамера размещается на расстоянии 50-60 см справа от испытуемого. Высота ее установки регулируется индивидуально для каждого испытуемого, который находится в положении сидя. Видеокамера должна размещаться в профиль испытуемого параллельно плоскости измерения.

2 Инициализация устройств. Настройка и калибровка спирометра производится в соответствии с руководством по его эксплуатации.

3 Сбор анамнеза (возраст, рост, вес, наличие заболеваний и в т.ч. хронических), ввод информации об испытуемом в базу спирометра. При этом в спирометре автоматически рассчитываются должные величины объемов дыхания (дыхательный объем, жизненная емкость легких и т.д.).

4 Подготовка к исследованию.

Испытуемого необходимо проинструктировать, как правильно выполнять спирометрические тесты, проконтролировать правильное положение головы, плеч и всего тела в целом (в положении сидя), спина испытуемого не должна соприкасаться со спинкой стула, руки не должны закрывать обзор крайних точек спины и грудной клетки (живота). Зубы должны лечь на мундштук, язык должен лежать на нижнем небе под мундштуком, губы плотно охватывают трубку. Перед выполнением теста необходимо перекрыть носовое дыхание носовым зажимом.

5 Выбор режима измерений.

Измерения производятся в режиме спирометрии, пневмотахометрии либо максимальной вентиляции легких [8]. В режиме спирометрии измерение жизненной емкости легких (ЖЕЛ) осуществляется на глубоком вдохе [8-9].

6 Запуск видеокамеры и спирометра, выполнение спирометрических тестов.

7 Запись данных спирометром и видеокамерой.

Спирометром определяется качество выполнения теста, при этом исследователь произносит команды согласно тесту и контролирует по кривой спирометрии и внешним признакам правильность выполнения команд и техники дыхания, а также наличие носового зажима, правильность положения тела, головы и рук [9]. В случае ошибок при

выполнении теста испытуемым, запись останавливается и начинается заново. Если количество попыток выполненных тестов превышает три-четыре измерения, то испытуемому необходимо дать отдых в течение 5-10 минут, и запись исследования осуществить повторно. По окончании теста печатается протокол (отчет) об исследовании.

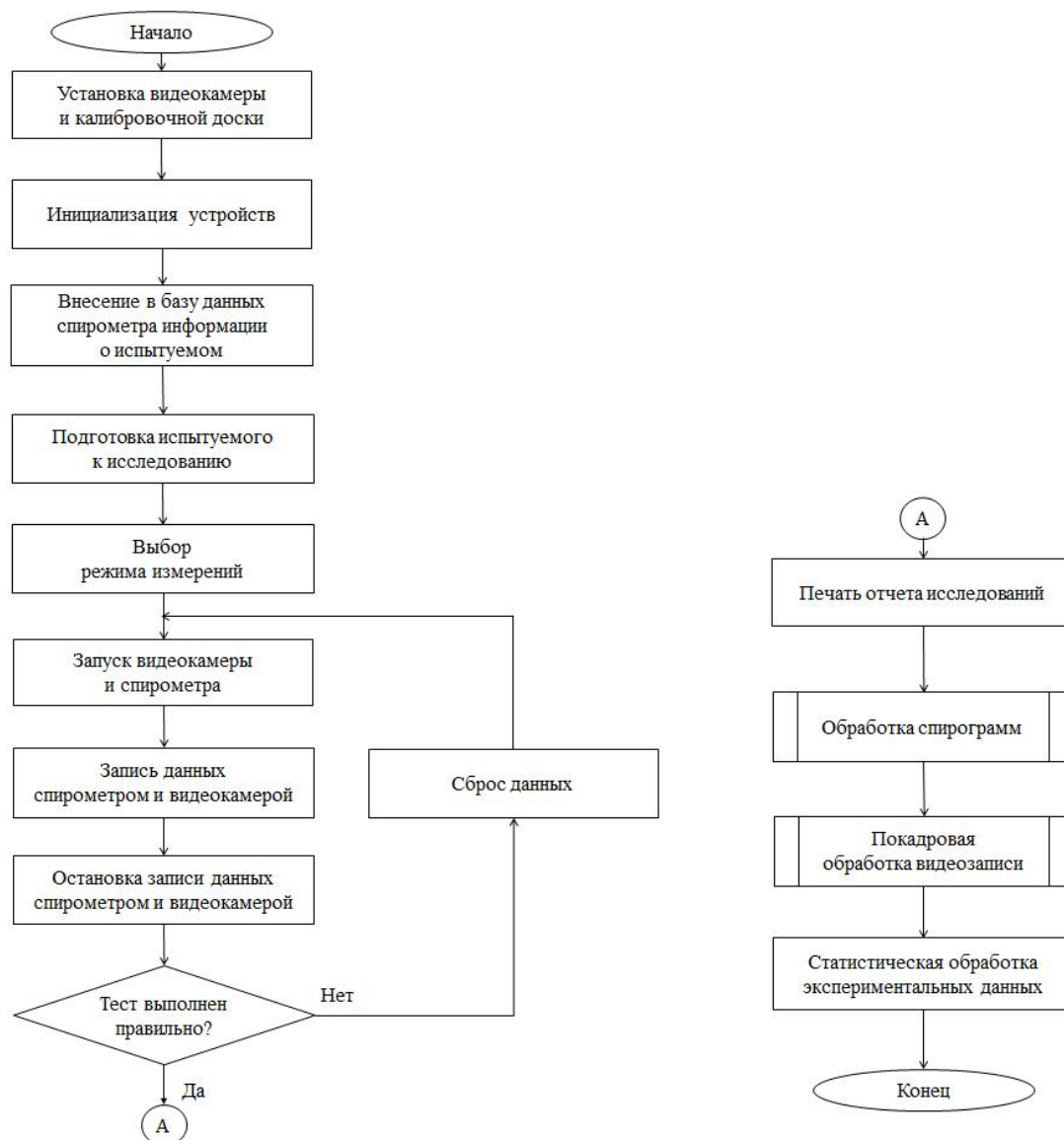


Рисунок 3. Алгоритм исследования и обработки данных биомеханики дыхания

8 Остановка записи данных.

9 Обработка спирограмм. Обработка данных включает этапы (рисунок 4):

- а) определение фаз спокойного и форсированного дыхания;
- б) определение точек экстремума по фазам дыхания;
- в) определение текущих объемов, соответствующих точкам экстремума;
- г) составление массива данных текущих объемов дыхания.

10 Покадровая обработка видеозаписи (рисунок 4):

- а) определение времени начала записи спирометрического теста;
- б) определение фаз спокойного и форсированного дыхания по видеозаписи;
- в) определение временных отсчетов, соответствующих найденным фазам дыхания;

г) извлечение видеофрагментов, соответствующих фазам дыхания. Каждый извлекаемый видеофрагмент сопоставляется с фиксируемыми точками на спирограмме, полученной со спирометра.

д) обработка полученных кадров. На каждом видеофрагменте по шести уровням определяются значения координат крайних точек спины x_{fi} (задней стенки) и передней стенки грудной клетки / живота x_{ai} .

е) расчет относительных отклонений стенок грудной клетки и живота осуществляется формуле:

$$\Delta x_i = x_{ai} - x_{fi}, \quad (1)$$

где x_{ai} – крайняя точка передней стенки грудной клетки/живота, определяемая по калибровочной доске;

x_{fi} – крайняя точка задней стенки (спины);

i – номер уровня отклонения. Уровни отклонений от 1 до 3 описывают грудное дыхание, уровни от 4 по 6 описывают брюшное дыхание.

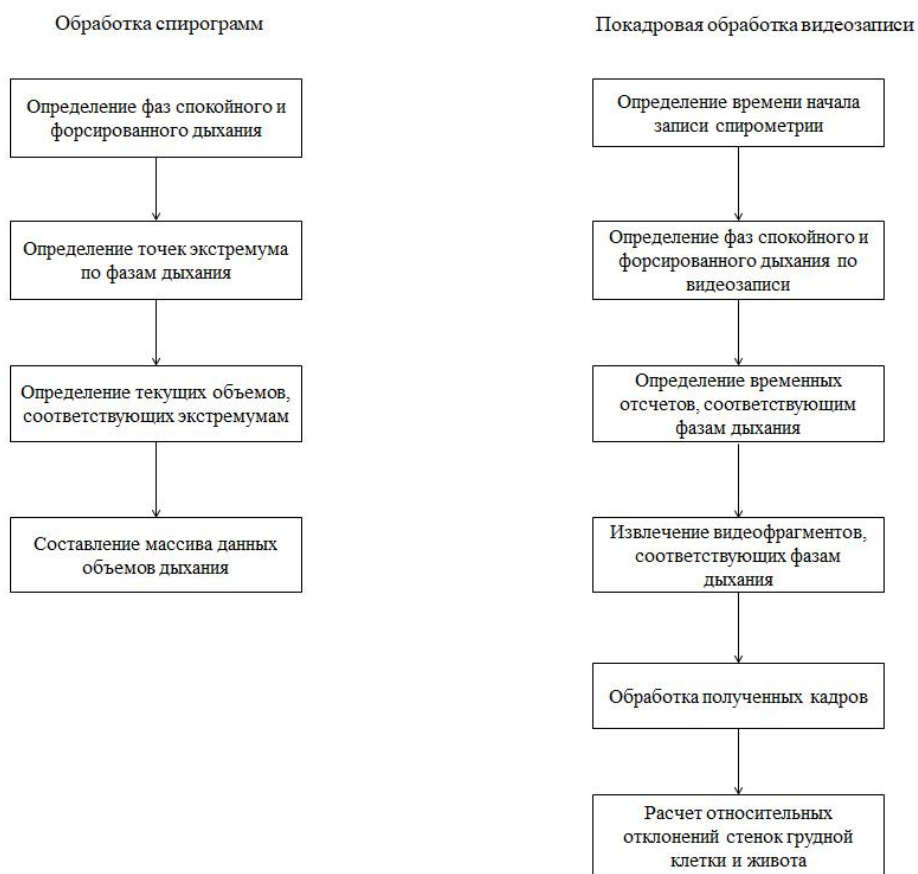


Рисунок 4. Алгоритм обработки спирограмм и покадровой обработки видеозаписи

Относительные отклонения стенок грудной клетки и живота определяли на калибровочной доске по шести уровням, равноудаленным по высоте друг от друга. Величины Δx_1 , Δx_2 , Δx_3 соответствуют уровням 1, 2, 3 и описывают грудное дыхание. Исходным уровнем являлась третья точка (уровень 3), которая на 3 см ниже уровня соска, по которой определяли величину отклонения Δx_3 . Величины Δx_1 и Δx_2 соответствуют

уровням на 8 и 16 см выше исходного уровня относительно калибровочной доски. Брюшное дыхание описывают отклонения Δx_4 , Δx_5 и Δx_6 , которые определялись ниже третьего уровня на 8, 16 и 24 см соответственно.

11 Статистическая обработка экспериментальных данных. Для оценки нелинейной связи между объемами вдыхаемого и выдыхаемого воздуха и значений отклонений Δx_i рассчитывается корреляционное отношение η по формуле:

$$\eta = \sqrt{1 - \frac{\sigma_{ост}^2}{\sigma^2}} = \sqrt{1 - \frac{(y_i - \bar{y})^2}{(y_i - \bar{y})^2}}, \quad (2)$$

где $\sigma_{ост}^2$ – остаточная дисперсия;

σ^2 – дисперсия фактических значений результативного признака;

y – теоретические значения, полученные по уравнению регрессии;

y_i – эмпирические значения;

Подкоренное выражение корреляционного отношения представляет собой коэффициент детерминации R^2 и рассчитывается по формуле:

$$R^2 = 1 - \frac{(y_i - \bar{y})^2}{(y_i - \bar{y})^2}. \quad (3)$$

Коэффициент детерминации R^2 показывает, насколько тесно наблюдаемые значения примыкают к линии регрессии. Для проверки адекватности регрессионной модели используется критерий Стьюдента t_η и рассчитывается по формуле:

$$t_\eta = \frac{\eta}{m_\eta} = \eta \sqrt{\frac{N-2}{1-\eta^2}}, \quad (4)$$

где $(N-2)$ – число степеней свободы при заданном уровне значимости и объеме выборки N ;

m_η – стандартная ошибка корреляционного отношения η .

Для корреляционного отношения рассчитывается стандартная ошибка, которую можно записать в виде: $\eta \pm m_\eta$.

Заключение. Разработанные аппаратно-программный комплекс и алгоритм исследований и оценки легочных объемов позволяет по значениям отклонений стенок грудной клетки и живота оценивать легочные объемы дыхания. Использование статистической обработки экспериментальных данных позволяет обрабатывать большие объемы данных.

Список литературы

[1] Ferkol, T.; Schraufnagel, D. The global burden of respiratory disease. Ann. Am. Thorac. Soc. 2014, 11, p. 404–406.

[2] Буцель, А. Ч. Острые респираторные вирусные инфекции: синдромальная диагностика, лечение и профилактика : учебно-методическое пособие / А. Ч. Буцель, Е. С. Яцкевич, Г. Г. Максименя. – Минск : БГМУ, 2020. – 48 с.

[3] Chu M., Nguyen T., Pandey V., Zhou Y., Pham H. N., Bar-Yoseph R., Radom-Aizik S. [et al.]. Respiration rate and volume measurements using wearable strain sensors. *npj Digital Medicine*, 2019, vol. 2, iss. 1, art. ID 8. <https://doi.org/10.1038/s41746-019-0083-3>.

[4] Moeuyersons, J.; Morales, J.; Seews, N.; Van Hoof, C.; Hermeling, E.; Groenendaal, W.; Willems, R.; Van Huffel, S.; Varon, C. Artefact Detection in Impedance Pneumography Signals: A Machine Learning Approach. *Sensors* 2021, 21, 2613. <https://doi.org/10.3390/s21082613>.

[5] Ревинская, И. И. Математическая модель для оценки объемных параметров дыхания = Mathematical Model for Estimation of Volume Respiratory Parameters / И. И. Ревинская // Доклады БГУИР. – 2023. – Т. 21, № 1. – С. 19-25.

[6] Revinskaya I. I., Kamlach P. V., Liashchevich E. I. Hardware-software complex for studying of breathing volume parameters. *Vestsi Natsyuanal'nai akademii navuk Belarusi. Seryya fizika-technichnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physical-technical series*, 2023, vol. 68, no. 2, pp. 149–155. <https://doi.org/10.29235/1561-8358-2023-68-2-149-155>.

[7] Ревинская, И. И. Методика оценки легочных объемов на основе записи биомеханики дыхательных движений = Methodology for Assessing Lung Volumes Based on Recording the Biomechanics of Respiratory Movements / И. И. Ревинская, П. В. Камлач // Доклады БГУИР. – 2023. – Т. 21, № 6. – С. 92–98.

[8] Дубровский, В. И. Биомеханика / В. И. Дубровский, В. Н. Федорова. М.: Владос-Пресс, 2003.

[9] Давыдченко, С. В. Исследование вентиляционной функции легких в клинической практике / С. В. Давыдченко, А. А. Бова. Минск: Белор. госуд. медицин. ун-т, 2007.

Авторский вклад

Ревинская Инна Ивановна – разработка аппаратно-программного комплекса, методики эксперимента, а также алгоритма обработки данных биомеханики дыхательных движений, подготовка рукописи статьи.

Камлач Павел Викторович – постановка целей и задачи исследования.

Лещевич Евгений Игоревич – разработка аппаратно-программного комплекса, методики эксперимента, подготовка рукописи статьи.

ALGORITHM FOR PROCESSING DATA OF BIOMECHANICS OF RESPIRATORY MOVEMENTS

I.I. Revinskaya
*Assistant at the Department of
Electronic Engineering and
Technology of BSUIR*

P.V. Kamlach
*Cand. of Sci., Associate
Professor, Associate Professor at
the Department of Electronic
Engineering and Technology*

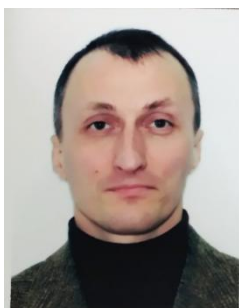
E.I. Liashchevich
*Postgraduate student of the
Department of Electronic
Engineering and Technology of
BSUIR*

Abstract. A hardware-software complex has been developed for studying the biomechanical movements of the walls of the chest and abdomen, as well as an algorithm for studying and processing data on the biomechanics of breathing. A methodology for processing research data is proposed.

Keywords. breathing, biomechanics, video camera, spirometry, lung volume, data processing algorithm.

УДК 004.67

МЕТОДИКА АНАЛИЗА БОЛЬШИХ МАССИВОВ ДАННЫХ ДЛЯ ОЦЕНКИ КАЧЕСТВА РАБОТЫ ВОДИТЕЛЯ И ТЕХНИЧЕСКОГО СОСТОЯНИЯ АВТОМОБИЛЯ



Д.Н. Одинаец
Доцент кафедры ЭВМ БГУИР,
кандидат технических наук,
доцент
adzinets@bsuir.by



Е.А. Алуев
Инженер-исследователь
АТЕК, бакалавр
технических наук,
alooeff@gmail.com

Д.Н. Одинаец

Окончил Ленинградский Военный Космический институт имени А.Ф.Можайского. Область научных интересов связана с распознаванием образов, моделированием сложных систем, обработкой больших массивов данных, параллельными вычислениями, сетевыми технологиями.

Е.А. Алуев

Выпускник Брестского государственного технического университета. Область научных интересов связана с облачными технологиями, машинным обучением и мульти-агентным моделированием.

Аннотация. Разработан метод сбора и анализа данных с шины CAN грузового автомобиля, на основе которого создана и протестирована методика интеллектуального анализа больших объемов данных на действующем автотранспортном предприятии.

Результатом работы методики является аналитическая оценка работы водителя и автомобиля (основные узлы и параметры движения) во время выполнения рейса. Показано, что созданная методика аналитической оценки работы водителя и транспортного средства (ТС) позволяет выявлять события по нарушению ПДД, конвенции ЕСТР (Европейского соглашения, касающегося работы экипажей транспортных средств, производящих международные автомобильные перевозки), правил эксплуатации и технического состояния ТС.

Ключевые слова: телеметрия, мониторинг автотранспорта, шина CAN, обработка больших объемов данных.

Введение. В настоящее время мировой объем автомобильных грузоперевозок неуклонно растет [1]. Выполнение этих грузоперевозок невозможно без использования водителей и транспортных средств.

Для обеспечения безопасности работы водителей в Женеве (Швейцария) 1 июля 1970 года было подписано Европейское соглашение, касающееся работы экипажей транспортных средств, производящих международные автомобильные перевозки (ЕСТР). Оно регламентирует режим труда и отдыха водителей, который контролируется дорожной полицией в странах, присоединившихся к нему. Вступило в силу для Республики Беларусь 02.10.1993 [2]. В настоящее время в этом списке 52 страны. Ответственность за нарушения положений конвенции ЕСТР и ПДД несет как водители, так и их предприятие-наниматель, поэтому все заинтересованы в минимизации риска нарушений.

Для минимизации таких материальных издержек эксплуатации транспортных средств, как расход топлива (затраты на него составляют около 40% всех затрат в грузоперевозках), износ тормозных систем и автошин предприятия стимулируют водителей разными методами, однако для эффективного анализа этих мер и результата их использования необходим комплексный подход в оценке качества работы водителя.

Для решения этих задач и с учетом достижений проекта *euroFOT* [3] была разработана методика анализа работы водителя и транспортного средства [4] на основании телеметрии [5], получаемой с автомобиля во время его работы. Практическое применение методики в разработанной системе [6] показало, что у разработанной системы есть потенциал расширения функций в части сбора большого количества данных о работе современного автомобиля с его шины *CAN* [7].

Целью данного исследования является разработка методики, которая позволит:

- Производить сбор сырых данных шины *CAN* автомобиля и передачу телеметрии на сервер для последующего хранения и обработки;
- Производить интеллектуальный анализ собранных данных.

Существуют работы по сбору некоторых данных о водителе и автомобиле при помощи смартфона [8], однако они имеют ряд таких серьезных недостатков, как низкая надежность, зависимость от электрического питания и низкая защита от вмешательства в их работу самим водителем.

Прототипом исследования является следующий проект: Система мониторинга грузового транспорта с использованием мониторингового блока *Teltonika FM4200* [9], которая позволяет собирать только некоторые данные с аналоговых и цифровых датчиков и ограниченный набор параметров шины *CAN*, передавать их на сервер хранения при помощи *GPRS*. Для работы блока с шиной *CAN* необходимо использовать блок-адаптер для каждой конкретной марки автомобиля для дешифровки передаваемых по шине пакетов данных.

Так как в текущем исследовательском проекте происходит сбор всех доступных пакетов информации, передаваемой по шине *CAN* и количество этих пакетов достигает 750 в секунду (например, в бортовом грузовике *KAMAZ* – около 250), то пропускной способности имеющегося радиоканала оказывается недостаточно. Для этого предлагается использовать мониторинговый блок *Вега MT X LTE* (Рисунок 1) с *LTE* и поддержкой до 3-х шин *CAN* одновременно, что дает возможность получать наиболее полную информацию о транспортном средстве.

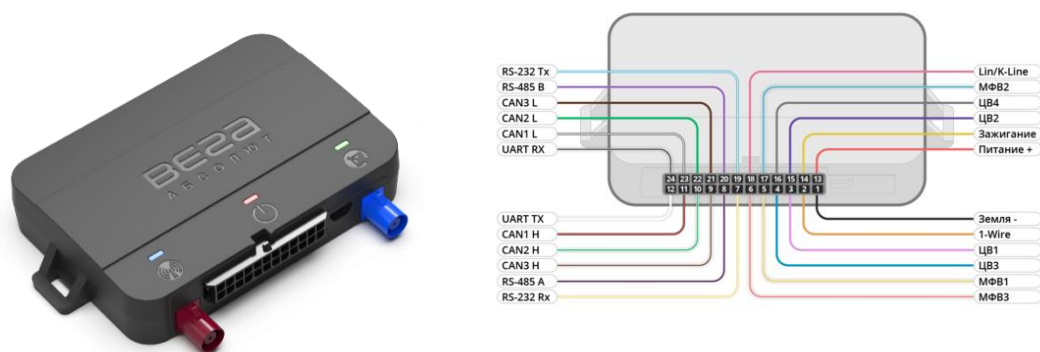


Рисунок 1. мониторинговый блок Вега *MT X LTE* – внешний вид и схема подключения

Как результат апробации блока на автомобиле *KAMAZ* за один месяц работы собирается около 8 GB информации (Пример на рисунке 2).

21,643	1	0005	0cf00400	8	F0 7D 7D E8 17 00 F3 7D 00000000
21,650	1	0005	18ffe03d	8	89 00 0D 7C 01 10 13 00 00000000
21,653	1	0005	0cf00400	8	F0 7D 7D D0 17 00 F3 7D 00000000
21,654	1	0005	0cf00300	8	C1 00 00 FF FF FF FF 00000000
21,660	1	0005	0cfe6cee	8	4B 10 C0 C0 14 13 80 17 00000000
21,661	1	0005	18fef11e	8	F3 FF FF C3 CC FF E0 F0 00000000
21,661	1	0005	18ffe03d	8	89 00 0D 7C 01 10 13 00 00000000
21,663	1	0005	0cf00400	8	F0 7D 7D A8 17 00 F3 7D 00000000
21,664	1	0005	18f0010b	8	C0 FF 30 FF FF 0D FF FF 00000000
21,670	1	0005	18ffe03d	8	89 00 0D 7C 01 10 13 00 00000000
21,673	1	0005	0cf00400	8	F0 7D 7D 90 17 00 F3 7D 00000000
21,675	1	0005	18f00f52	8	5C 1D F1 ED 55 1F FF FF 00000000
21,680	1	0005	18ffe03d	8	89 00 0D 7C 01 10 13 00 00000000
21,682	1	0005	18febf0b	8	7D 17 7C 7D 79 7A FF FF 00000000
21,683	1	0005	0cf00400	8	F0 7D 7D 68 17 00 F3 7D 00000000

Рисунок 2. Пример части данных, собираемых с шины CAN автомобиля KAMAZ

Хранение данных. В некоторых системах количество собираемой информации настолько велико, что ее получение и передача во время движения автомобиля невозможна в полном объеме из-за аппаратных ограничений мониторинговых блоков и каналов связи. В реализуемом проекте производительности мониторингового блока Вега MT X LTE оказалась достаточной для сбора данных с автомобиля KAMAZ и передачи их в хранилище AWS S3 в формате Apache Parquet при помощи сервиса AWS IoT FleetWise.

Дешифровка пакетов шины CAN. В связи с тем, что каждый автопроизводитель кодирует данные о агрегатах автомобиля в пакетах по своему алгоритму, необходима процедура декодирования. Ключи для декодирования официально не распространяются, однако доступны для большинства распространенных марок автомобилей.

Предварительный анализ данных. Доступ к данным и их обработка осуществляется на платформе Nadoor в соответствии с алгоритмом на рисунке 3. На первом этапе производится предварительная обработка собранных данных (профайлинг) с датчиков и CAN-шины с целью оценки их качества, и при необходимости применения к ним средств фильтрации данных: заполнение пропусков, подавление аномальных значений, исключение дубликатов и противоречий. Для устранения погрешности измерений, которые представляют собой «шум» в данных, производится анализ и устранение заведомо некачественных данных (выход значений за допустимые пределы, пропуски в данных, резкие нереальные изменения некоторых параметров, например, из-за «дребезга» контактов механических датчиков).

Работа с наблюдаемыми и расчетными параметрами. Некоторые параметры предварительно обрабатываются для получения расчетных параметров на основании вариаций вектора изменения наблюдаемых параметров. Например, скорость движения, показания акселерометра позволяют вычислить прямое и поперечное ускорения, а обороты двигателя, скорость ТС и номер передачи КПП позволяет вычислить нагрузку на двигатель. Затем полученные расчетные параметры используются совместно с первоначальными параметрами на следующих этапах обработки (Рисунок 3).

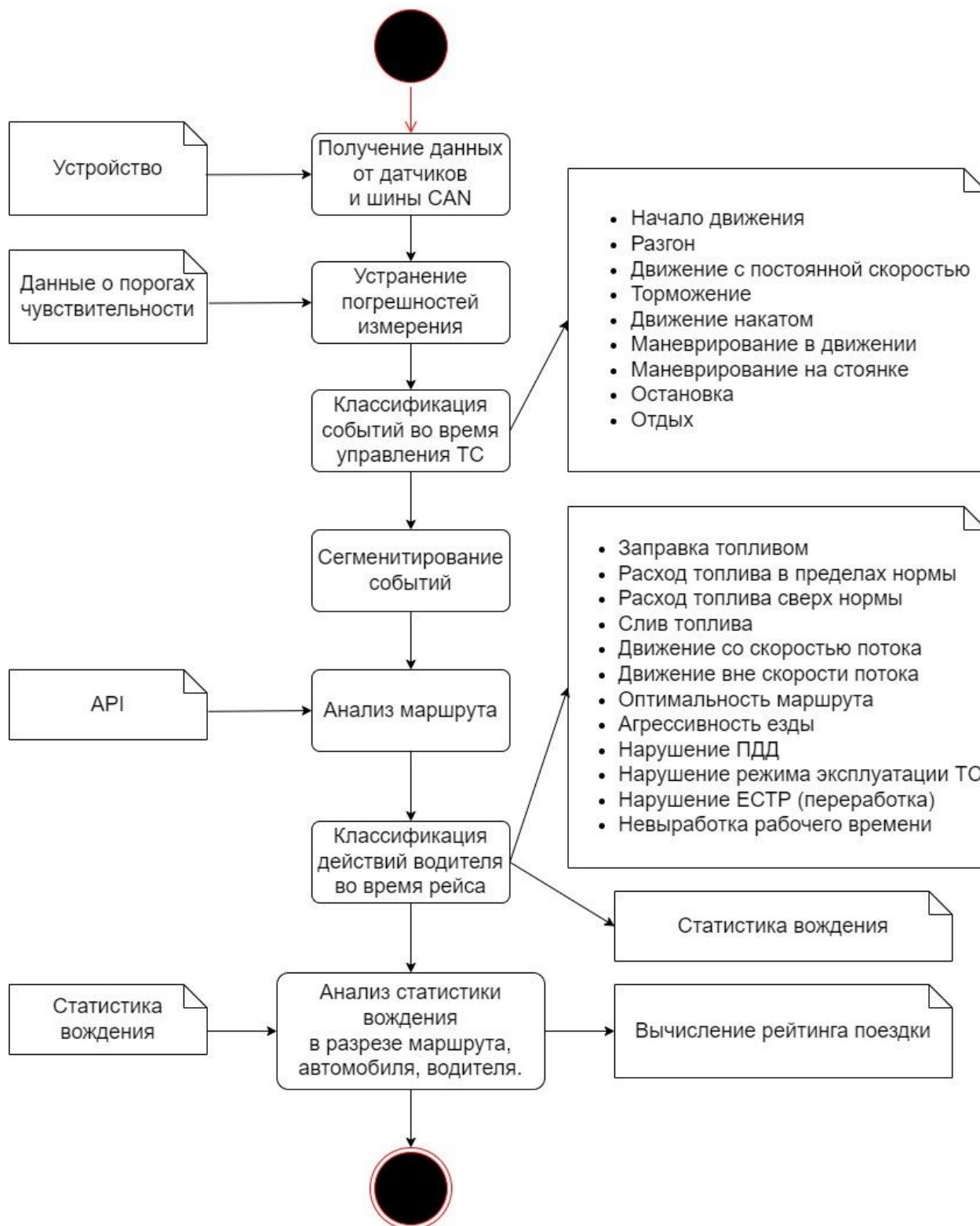


Рисунок 3. Алгоритм обработки и анализа данных

Сегментирование событий. Для упрощения обработки данных маршрут движения разбивается на сегменты, которые обрабатываются по отдельности для классификации событий во время управления ТС, а затем как единое целое – для классификации действий водителя по время рабочей смены.

Выявление и классификация событий во время управления ТС. Производится классификация таких событий поездки как начало движения, разгон, движение с постоянной скоростью, торможение, движение накатом, маневрирование в движении, маневрирование на стоянке, остановка и отдых.

Например, событие «Движение с постоянной скоростью» определяется как поддержание скорости в пределах ± 10 км/ч (изменение скорости более 10 км/ч рассматриваем как изменение дорожной обстановки). Критерий равномерности скорости движения при этом событии рассчитывается по количеству циклов изменения вектора скорости за единицу времени. Колебания скорости до 2 км/ч вызваны гистерезисом круиз контроля и не учитываются системой. При возникновении замедления впереди идущего автомобиля несоблюдение дистанции вынуждает производить торможение, в результате появляется волновой эффект пульсаций скорости всего потока машин.

Событие «Разгон» имеет критерий «Разгон», который показывает эффективность разгона, который должен происходить в зеленой зоне тахометра. Если разгон слишком медленный, то при наличии АКПП переключение на следующую передачу приводит к падению оборотов до 850 об/мин, а зеленый сектор эффективной работы турбины начинается с 1000 об/мин. Если же разгон слишком интенсивен, то АКПП переключает передачи в диапазоне 1300-1650 об/мин, что также выходит за пределы зеленого сектора тахометра. При любом типе КПП слишком интенсивный разгон приводит к большим нагрузкам в системе привода и шинах.

Для каждого события в системе задан вес, который при анализе данных, вместе с продолжительностью события, влияет на результат критерия.

Анализ маршрута. Производится анализ выбранного водителем маршрута на предмет дорожной обстановки (транспортные заторы, дорожные ограничения) для анализа, насколько скорость движения ТС соответствует разрешенной на отдельных участках пути с использованием данных *WebAPI* от ООО «МИТ» [10].

Классификация действий водителя во время работы. Производится классификация действий водителя, связанных с выполнением рейса:

- Заправка топливом;
- Расход топлива в пределах нормы;
- Расход топлива сверх нормы;
- Слив топлива;
- Движение со скоростью потока;
- Движение вне скорости потока;
- Агрессивность езды – выход за установленные пределы значений ускорений (положительное для разгона, отрицательное для торможения и боковое для маневрирования);
- Нарушение ПДД – превышение максимально разрешенной скорости движения, превышение максимально допустимой нагрузки на ось, движение во время действия ограничений (весенние, из-за паводков и летние, из-за высокой температуры), движение в зонах ограничений движения для большегрузных ТС;
- Нарушение режима эксплуатации ТС – анализ таких параметров, как обороты двигателя, давление в шинах, амплитуду вибрации и фиксация случаев выхода этих параметров за заданные границы;
- Нарушение технического состояния ТС – выявление превышения мгновенного расхода топлива над расчетным для текущей скорости, оборотов двигателя и величины ускорения, а также ошибок, связанных с выходом из строя компонентов ТС;
- Нарушение конвенции ЕСТР – отсутствие ежедневного периода отдыха, превышение ежедневной продолжительности управления, допускаемой по конвенции ЕСТР;

– Невыработка рабочего времени – ежедневная продолжительности управления сильно меньше 9 часов приводит к неиспользованию рабочего времени для движения (требует ручного анализа, так как может быть связана с временными ограничениями на дорогах или ожиданием готовности груза к перевозке, либо очередями при прохождении пограничного контроля).

Анализ статистики вождения. На основании данных за текущий и предыдущие рейсы производится анализ статистики вождения в разрезе маршрута, автомобиля, водителя. Результатом анализа являются вычисленные рейтинг поездки и стиль вождения водителя.

Система имеет возможность анализировать статистику как в пределах работы одного водителя, так и нескольких. В первом случае удастся оценить изменение стиля на шкале времени в зависимости от маршрута. Во втором случае можно сопоставить стили вождения разных водителей на одних и тех же маршрутах, но принимая во внимание изменение дорожной обстановки (заторы, сезонные ограничения).

Применение на практике. Описанный способ сбора и анализа данных о работе водителей и транспортных средств была применена на практике [6]. Результатом запуска системы стал больший объем данных для анализа. В результате были выявлены ряд кадровых, технических и организационных закономерностей. Они были проанализированы и по результатам были приняты решения, позволяющие улучшить работу автотранспортного предприятия как в краткосрочной, так и в долгосрочной перспективе.

Полученные предприятием данные также были проанализированы для внесения доработок и исправлений в аппаратно-программную систему, что позволит повысить как эффективность ее работы, так и качество и оперативность получаемых данных.

Использование современных архитектурных решений позволило сократить затраты по подключению системы к транспортным средствам разных марок и доступу к серверам хранения информации.

Заключение. Разработанная методика анализа оценки работы водителя и транспортного средства на основе данных с шины *CAN* оперирует большими объемами данных и позволяет выявлять и классифицировать гораздо больше данных, получаемых от транспортного средства, чем в предыдущем проекте [4]. В дополнение к тактическим событиям непосредственного управления ТС и стратегическим событиям о поведении на дороге и выполнения работы водителем во время рейса получена возможность детального контроля состояния узлов и агрегатов автомобиля во время движения. Полнота данных о состоянии автомобиля напрямую зависит от доступности информации по кодировке пакетов шины *CAN* и, соответственно, от возможности декодирования этих пакетов. Также стоит отметить, что современные грузовые автопоезда с пневмоподвеской *Wabco* позволяют отследить не только состояние дорожного полотна, развесовку груза по осям, но и безопасность стиля вождения автомобилем. Например, во время одного рейса был дважды зафиксирован отрыв всех колес правой стороны полуприцепа от дорожного полотна, что было вызвано резким маневрированием водителя, не соблюдающим дистанцию до впереди движущегося автомобиля при неблагоприятных погодных условиях. Своевременное выявление подобных инцидентов позволяет заблаговременно провести необходимые инструктажи по технике безопасности и безопасному вождению.

Для реализации эффективной обработки и анализа собранных с автомобилями больших объемов данных в рамках предложенной методики задействованы современные возможности облачной платформы *AWS*, а именно сервиса *Elastic MapReduce* с использованием экосистемы *Hadoop* на основе фреймворка *MapReduce*. Гибкость предложенной методики позволяет оперативно вносить изменения в алгоритмы

классификации действий водителя, анализа маршрута и накопленной статистики, однако требует высокой квалификации специалистов.

Список литературы

- [1] UNECE global SDG indicators <https://w3.unece.org/SDG/en/Indicator?id=88>
- [2] ПОСТАНОВЛЕНИЕ СОВЕТА МИНИСТРОВ РЕСПУБЛИКИ БЕЛАРУСЬ от 30 ноября 1992 г. № 721 О присоединении Республики Беларусь к международным договорам, регламентирующим перевозку грузов в международном автомобильном сообщении <https://etalonline.by/document/?regnum=c29200721>
- [3] Mohamed Benmimoun, Felix Fahrenkrog, Dr. Adrian Zlocki, Prof. Dr.-Ing. Lutz Eckstein. INCIDENT DETECTION BASED ON VEHICLE CAN-DATA WITHIN THE LARGE SCALE FIELD OPERATIONAL TEST “euroFOT”. Institut für Kraftfahrzeuge, RWTH Aachen University
- [4] Eugene Alooeff ML Based Methodology of The Truck Driving Evaluation ICMCE 2024
- [5] ГОСТ 19619-74 Оборудование радиотелеметрическое. Термины и определения. — 1975
- [6] АТЕК <https://atek.dev/research>
- [7] ISO 11898-1:2015 Road vehicles. Controller area network (CAN) <https://www.iso.org/standard/63648.html>
- [8] Лашков И.Б. Анализ поведения водителя при управлении транспортным средством с использованием фронтальной камеры смартфона. ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ 4(89)6 2017 с.7-17
- [9] Руководство пользователя FM4200 <https://www.euromobile.ru/upload/iblock/5c0/5c0e969470cfb9983dc220e54aabab50.pdf>
- [10] WebAPI Homepage. <https://probki.net/b2b/webapi.html>

Авторский вклад

Одинец Дмитрий Николаевич – руководство проектом по разработке методики сбора и анализа данных шины CAN грузового автомобиля, анализ текущей ситуации на рынке мониторинговых систем.

Алуев Евгений Александрович – разработка архитектуры системы, техническая реализация проекта, анализ полученных результатов.

A TECHNIQUE FOR BIG DATA ANALYSIS TO ASSESS THE QUALITY OF THE DRIVER'S WORK AND THE TECHNICAL CONDITION OF THE TRUCK

Adzinets D.

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

Alooeff E.

R&D Engineer ATEK, Bachelor of Computer Science, Bachelor.

Abstract. A method for collecting and analyzing data from the truck CAN bus has been developed, on the basis of which a system for intelligent analysis of large volumes of data has been created and tested at an operating transport enterprise.

The result of the system's operation is analytical assessment of the truck driver's work and the vehicle operation (main components and movement parameters) during the road trip. It is shown that the created method of analytical assessment of the driver work and the vehicle makes it possible to identify events related to violation of traffic rules, the AETR convention (European Agreement concerning the Work of Crews of Vehicles Engaged in International Road Transport), operating rules and technical condition of the vehicle.

Keywords: Telemetry, truck monitoring, CAN bus, big data processing.

УДК 004.021:004.832.2

ПРЕИМУЩЕСТВА И НЕДОСТАТКИ ЭМПИРИЧЕСКИХ МОДЕЛЕЙ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ В ОБЛАСТИ ЭЛЕКТРОННОЙ КОММЕРЦИИ



Михальков М.Д.

*магистрант кафедры программного обеспечения информационных технологий факультета компьютерных систем и сетей БГУИР
solution.m@mail.ru*

М.Д. Михальков

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с анализом и разработкой моделей и алгоритмов построения интеллектуальных систем поддержки принятия решений.

Аннотация. Данная статья посвящена рассмотрению интеллектуальных систем поддержки принятия решений. Произведен анализ основных эмпирических моделей интеллектуальных систем поддержки принятия решений, таких как модели, основанные на нейросетях, модели на основе метода опорных векторов, модели, применяющие генетические алгоритмы. Выявлены преимущества и недостатки каждой из моделей. На основе произведенного анализа моделей отмечены наиболее эффективные в задачах из области электронной коммерции.

Ключевые слова: Интеллектуальные системы поддержки принятия решений, эмпирические модели, нейронные сети, метод опорных векторов, генетические алгоритмы.

Введение. Под термином «поддержка принятия решений» понимается совокупность процедур (обработка данных измерений, их анализ, формирование списка альтернатив, выбор лучшего решения и т. д.), обеспечивающих лицо, принимающее решения, необходимой информацией и рекомендациями, облегчающими процесс принятия решений [1].

Интеллектуальная система поддержки принятия решений представляет собой программно-аппаратный комплекс, использующий и интегрирующий элементы технологий искусственного интеллекта и информационных систем для обработки значительных объемов объективной и субъективной информации (эмпирические модели), а также теории игр и классического математического моделирования (аналитические модели).

В основу архитектуры интеллектуальной системы поддержки принятия решений положены следующие интеллектуальные способности лица, принимающего решения, и высококвалифицированных специалистов:

– выявлять скрытые закономерности в данных, позволяющих связывать различные факты одного события, относить объект к одному из заранее известных классов, прогнозировать будущие значения величин на основании их фактических значений,

разбивать совокупность объектов на однородные группы и осуществлять поиск существующих структур в данных;

– принимать обоснованные решения, то есть осуществлять поиск, сбор, представление информации о проблемной области в удобном для анализа виде, разработку альтернатив и выбор рационального варианта решения, доведение его до исполнителей, контроль и внесение необходимых корректив;

– обучаться, формировать новый опыт решения задачи в результате использования накопленного опыта решения аналогичных задач;

– адаптироваться к изменяющимся условиям в процессе решения задачи: корректировать очередную попытку решения задачи и осуществлять поиск новых методов ее решения;

– обрабатывать нечеткую, противоречивую, недостаточную или избыточную информацию [2].

В данной статье речь пойдет об эмпирических моделях интеллектуальных систем поддержки принятия решений: модели, основанные на нейросетях, модели на основе метода опорных векторов, модели, применяющие генетические алгоритмы. На основе анализа каждой из моделей будут выявлены их преимущества и недостатки с целью определения эффективности их использования в задачах, связанных с областью электронной коммерции.

Эмпирические модели интеллектуальных систем поддержки принятия решений на основе метода опорных векторов. Метод опорных векторов – это набор схожих алгоритмов вида «обучение с учителем», использующихся для задач классификации и регрессионного анализа [3]. Этот метод относится к линейным классификаторам.

Основная идея метода опорных векторов заключается в переводе исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве.

Две параллельных гиперплоскости строятся по обеим сторонам гиперплоскости, разделяющей классы. Разделяющей гиперплоскостью будет гиперплоскость, максимизирующая расстояние до двух параллельных гиперплоскостей. Чем больше расстояние между этими параллельными гиперплоскостями, тем меньше будет каждый объект данных представлен как вектор (точка) в пространстве (последовательность чисел). Это типичный случай линейной разделимости. Гиперплоскостей может быть довольно много, но необходимо найти такую гиперплоскость, чтобы расстояние от нее до ближайшей точки было максимальным. Расстояние между двумя ближайшими точками, располагающимися по разные стороны гиперплоскости, является максимальным. Существование такой гиперплоскости вызывает наибольший интерес, ее называют оптимальная разделяющая гиперплоскость, а соответствующий ей линейный классификатор – оптимальный разделяющий классификатор [4].

Пример разделяющей гиперплоскости АВ для некоторого множества точек представлен на рисунке 1. Ближайшие к параллельным гиперплоскостям точки называются опорными векторами.

Проблема построения оптимальной разделяющей гиперплоскости сводится к минимизации перпендикуляра к разделяющей гиперплоскости. Это задача квадратичной оптимизации, методы решения которых проработаны. Среди них можно выделить, последовательное квадратичное программирование, метод Ньютона, метод сопряженных направлений.

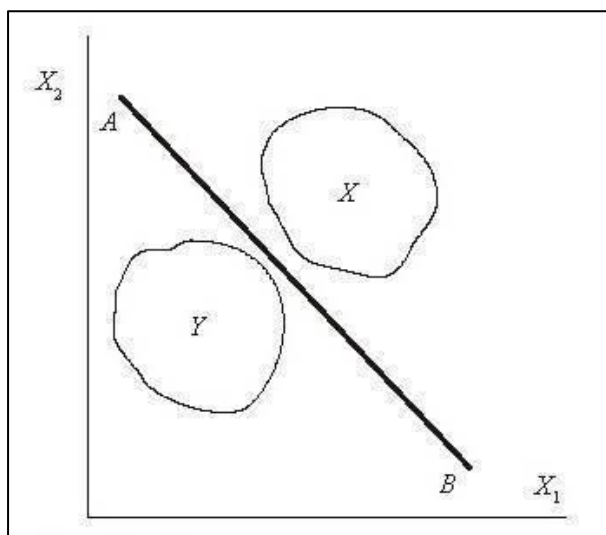


Рисунок 1. Пример разделяющей гиперплоскости АВ.

Основной идеей метода опорных векторов является перевод исходных векторов в пространство более высокой размерности, что накладывает некоторые ограничения на максимально допустимую размерность вектора исходных данных. Данная идея берет свое начало из предположений, что не все исследуемые множества данных являются линейно разделимыми, что требуется для применения метода разделения гиперплоскостями. При изначально высокой размерности исходного вектора перевод его в еще более высокоразмерное пространство ведет к усложнению алгоритмов анализа данной системы и требует больших вычислительных мощностей.

Подводя итог по методу опорных векторов, можно выделить следующие преимущества моделей на его основе:

- хорошо работает с пространством признаков большого размера;
- хорошо работает с данными небольшого объема;
- алгоритм максимизирует разделяющую полосу, которая позволяет уменьшить количество ошибок классификации;
- так как алгоритм сводится к решению задачи квадратичного программирования в выпуклой области, то такая задача всегда имеет единственное решение (разделяющая гиперплоскость с определенными гиперпараметрами алгоритма всегда одна).

К недостаткам моделей, основанных на методе опорных векторов, можно отнести следующие:

- долгое время обучения для больших наборов данных;
- неустойчивость к шуму: выбросы в обучающих данных становятся опорными объектами-нарушителями и напрямую влияют на построение разделяющей гиперплоскости;
- не описаны общие методы построения ядер и спрямляющих пространств, наиболее подходящих для конкретной задачи в случае линейной неразделимости классов.

Эмпирические модели интеллектуальных систем поддержки принятия решений, основанные на нейросетях. Искусственные нейронные сети – математические модели, и их программные или аппаратные реализации, построенные по принципу организации и функционирования биологических нейронных сетей (сетей нервных клеток живого организма, в том числе человека) [5].

Искусственные нейронные сети представляют собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов). Такие процессоры обычно довольно просты. Каждый процессор подобной сети имеет дело

только с сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим процессорам.

Важнейшее свойство нейронных сетей, свидетельствующее об их огромном потенциале и широких прикладных возможностях, состоит в параллельной обработке информации одновременно всеми нейронами. Благодаря этой способности при большом количестве межнейронных связей достигается значительное ускорение процесса обработки информации. Во многих ситуациях становится возможной обработка сигналов в режиме реального времени. Очень большое количество межнейронных соединений приводит к тому, что сеть становится нечувствительной к ошибкам, возникающим в отдельных контактах. Функции поврежденных соединений принимают на себя другие элементы, в результате в деятельности сети не наблюдаются заметные нарушения. Это свойство используется, в частности, при поиске оптимальной архитектуры нейронной сети путем разрыва отдельных связей.

Другое не менее важное свойство нейронной сети состоит в способности к обучению и к обобщению полученных знаний. Сеть обладает чертами так называемого искусственного интеллекта. Натренированная на ограниченном множестве обучающих выборок, она обобщает накопленную информацию и вырабатывает ожидаемую реакцию применительно к данным, не обрабатывавшимся в процессе обучения.

При решении задач прогнозирования роль нейронной сети состоит в предсказании будущей реакции системы по ее предшествующему поведению [6].

Рассмотрим методику построения модели на основе нейронной сети. В качестве модели искусственной нейронной сети с прямой связью, отображающей набор входных данных в набор выходных данных, был выбран многослойный персептрон (MLP).

Многослойный персептрон состоит из линейной функции активации во всех нейронах, то есть имеет простой механизм включения-выключения, позволяющий определить, сработает ли нейрон. Отличительной особенностью MLP является то, что каждый нейрон использует нелинейную функцию активации. Это позволяет смоделировать механизм действия биологических нейронов мозга. Данная функция имеет множество видов с единственным требованием – должна быть нормализуемой и дифференцируемой.

Многослойный персептрон состоит из одного входного и одного выходного слоя с одним или несколькими скрытыми слоями нелинейно-активирующих узлов. Каждый узел в одном слое соединяется с определенным весом для каждого другого узла в следующем слое. Система формирует модель на основе построения алгоритма, основанную на зависимости значений одних показателей от определенного ряда других параметров.

Как и любая нейронная сеть, построенная модель требует обучения (в нашем случае алгоритм обучения применялся к динамически построенной нейронной сети для минимизации ошибки путем непрерывного обучения сети до достижения оптимального уровня эффективности). Обучение происходит в персептроне путем изменения весов соединений (или синаптических весов) после обработки каждого фрагмента данных на основе количества ошибок в выходных данных по сравнению с ожидаемым результатом.

Эффективность выполнения подхода проверяется путем проведения сравнительного исследования, которое включает в себя тестирование динамически построенной сети и представляет сравнительный анализ результатов классификации.

Пример многослойной нейронной сети можно наблюдать на рисунке 2.

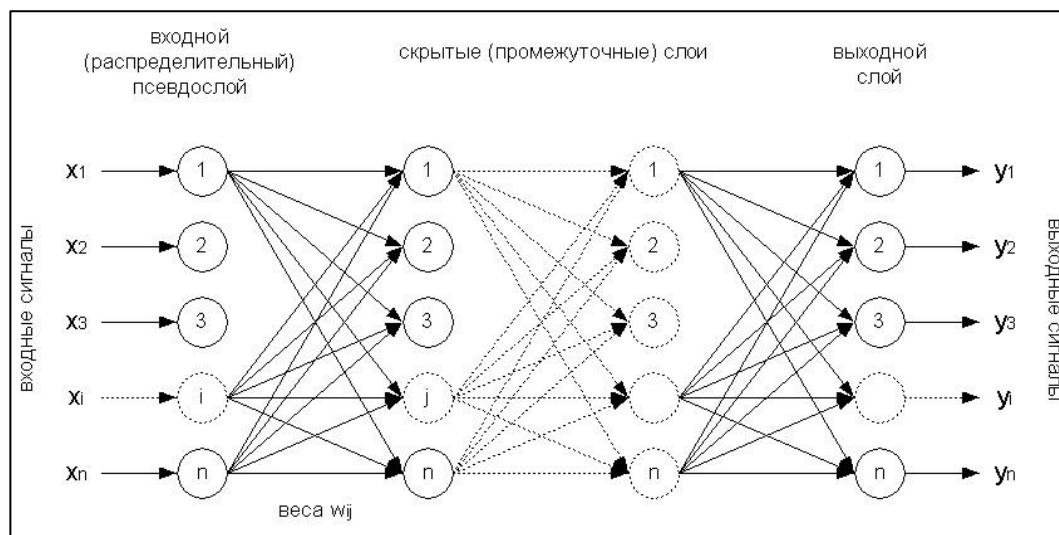


Рисунок 2. Пример многослойной нейронной сети

К преимуществам данной модели относятся:

- обучение без ограничений по сложности на сверхбольших объемах данных;
- возможность обучения без учителя.

Недостатками являются:

- для обучения требуются довольно большие объемы данных;
- для работы необходима мощная аппаратная база.

Эмпирические модели интеллектуальных систем поддержки принятия решений, применяющие генетические алгоритмы. Генетический алгоритм был предложен более 20 лет назад Джоном Холландом. Генетический алгоритм реализует метод случайного поиска по аналогии с теорией эволюции Дарвина. Основан на естественном отборе – основном механизме эволюции, работающем по принципу «выживает наиболее приспособленный». Подражая этому процессу, генетический алгоритм дает возможность находить близкие к оптимальным решения задачи [7].

Алгоритм управляет набором представителей, которые могут рассматриваться как возможные решения поставленной задачи. Отобранные на каждом шаге лучшие представители перемешиваются («скрещиваются») между собой, производя «мутантов». От полученных таким образом новых представителей ожидают еще более хороших результатов. Процесс повторяется несколько раз пока не будет достигнуто целевое значение [8].

Пошагово алгоритм выглядит следующим образом.

Шаг 1. Конструирование начальной популяции определенного размера (оператор инициализации).

Шаг 2. Выбор двух родительских хромосом из популяции (оператор селекции).

Шаг 3. Копирование выбранных хромосом и применение генетических операторов для создания новых хромосом (операторы кроссинговера, мутации).

Шаг 4. Отбор и последующее удаление хромосом из популяции для восстановления ее первоначального размера.

Шаг 5. До тех пор, пока не пройдено заданное число шагов, возвращаемся к шагу 2, иначе оканчиваем работу алгоритма.

Графически алгоритм можно представить так, как показано на рисунке 3.

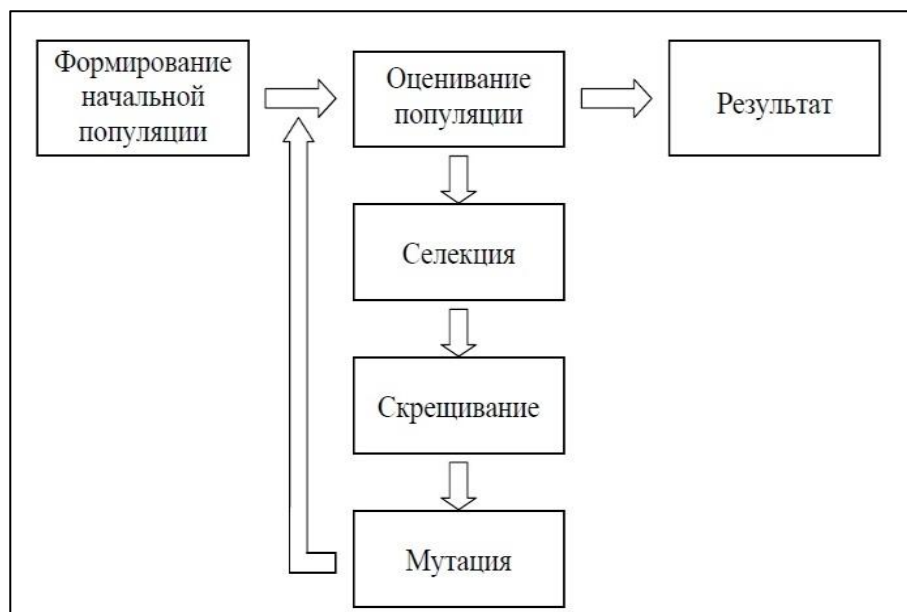


Рисунок 3. Генетический алгоритм

Достоинства применения алгоритма:

- применяется для решения сложных неформализованных задач, для которых не разработано специальных методов;
- имеет преимущества перед другими алгоритмами при очень больших размерах задач и отсутствия упорядоченности в исходных данных;
- алгоритм выполняется существенно быстрее других алгоритмов поиска на большом пространстве значений и значительно экономит память компьютера.

Недостатки алгоритма:

- генетический алгоритм не гарантирует, что найденное решение будет оптимальным (приемлем для поиска «достаточно хорошего» решения задачи за «достаточно короткое время»);
- в случаях, когда задача может быть решена специально разработанным для неё методом, практически всегда такие методы будут эффективнее генетического алгоритма как по быстрдействию, так и по точности найденных решений;
- неэффективен на небольшом пространстве поиска.

Заключение. На основе проведенного анализа основных эмпирических моделей интеллектуальных систем поддержки принятия решений можно сделать следующий вывод об эффективности их использования в задачах, связанных с областью коммерции.

Первая модель (на основе опорных векторов), несмотря на свои очевидные преимущества, будь то наглядность процедуры классификации, или простота в использовании, соответственно, не подходит для решения задач из области электронной коммерции, в первую очередь из-за своих ограниченных возможностей по работе с большими объемами входной выборки (многокритериальность, количество входных векторов).

Модели на основе нейросетевой классификации и генетических алгоритмах видятся более подходящими для решения подобного подмножества задач. Оба эти метода сводятся к решению задач оптимизации, для решения которых имеется большое количество численных и итерационных методов. Кроме того, нейронные сети решают проблему анализа линейно неразделимых пространств увеличением количества слоев сети.

Соответственно, для решения задач в области электронной коммерции среди основных эмпирических моделей наибольший интерес представляют именно модели, основанные на нейросетях.

Список литературы

- [1] Лазарсон Э.В. Формализация знаний и интеллектуальная поддержка принятия решений в задачах выбора // Интеллектуальные системы в производстве. 2006. № 2. С. 4-14.
- [2] Аксенов, К.А. Моделирование и принятие решений в организационно-технических системах: учебное пособие. В 2 ч. Ч.1 / К.А. Аксенов, Н.В. Гончарова. – Екатеринбург: Изд-во Урал. ун-та, 2015. – 104 с.
- [3] Ларин Р.М., Плясунов А.В., Пяткин А.В. Методы оптимизации. Примеры и задачи: Учеб. пособие. Новосибирск, 2003.
- [4] Burges C.J.C. A Tutorial on Support Vector Machines for Pattern Recognition., Boston, 1998.
- [5] Зайченко Ю.П. Нечеткие модели и методы в интеллектуальных системах; под общ. ред. М.З. Згуровского. Киев: Изд. дом «Слово», 2008. 333 с.
- [6] Авдеев А.С. Разработка программного комплекса нейропрогнозирования [Текст]/ А.С. Авдеев, О.И. Пятковский // Программные продукты и системы. – Тверь, 2010. – №1. С. 106-109.
- [7] Панченко Т.В. Генетические алгоритмы: учебно-методическое пособие / под ред. Ю.Ю. Тарасевича. Астрахань: Издательский дом «Астраханский университет», 2007. 87 с.
- [8] Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы / Пер. с польск. И.Д. Рудинского. М.: Горячая линия - Телеком, 2006. 452 с.

Авторский вклад

Михальков Максим Дмитриевич – анализ эмпирических моделей интеллектуальных систем поддержки принятия решений с целью определения их эффективности при использовании в задачах из области электронной коммерции.

ADVANTAGES AND DISADVANTAGES OF EMPIRICAL MODELS OF INTELLIGENT DECISION SUPPORT SYSTEMS IN THE FIELD OF E-COMMERCE

M.D. Mikhalkov

*PhD student at the Department of
Information Technology Software,
Faculty of Computer Systems and
Networks, BSUIR*

Abstract. This article is devoted to the consideration of intelligent decision support systems. An analysis of the main empirical models of intelligent decision support systems, such as models based on neural networks, models based on the support vector machine, and models using genetic algorithms, is carried out. The advantages and disadvantages of each model are revealed. Based on the analysis of the models, the most effective ones in issues from the field of e-commerce were noted.

Keywords: Intelligent decision support systems, empirical models, neural networks, support vector machines, genetic algorithms.

УДК 004.94

ОБРАБОТКА БОЛЬШИХ ДАННЫХ ПРИ ИССЛЕДОВАНИИ ВЗАИМОДЕЙСТВИЯ ИНФРАЗВУКА С БИОЛОГИЧЕСКИМИ ТКАНЯМИ



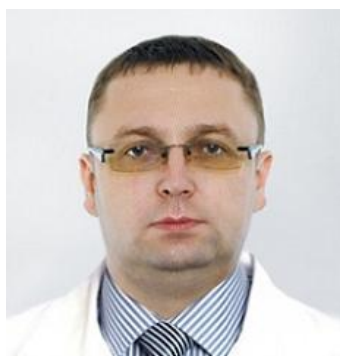
Е.И. Лецевич
аспирант БГУИР, инженер
кафедры электронной техники и
технологии БГУИР



П.В. Камлач
заместитель декана
факультета
компьютерного
проектирования, кандидат
технических наук, доцент



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



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



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



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

Е.И. Лецевич

Окончил БГУИР. Аспирант кафедры электронной техники и технологии. Область научного интереса – влияние инфразвука на биологические ткани.

П.В. Камлач

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

А.В. Чураков

Доцент кафедры электронной техники и технологии, кан.мед.наук, врач анестезиолог-реаниматолог

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 БГУИР. Ассистент кафедры электронной техники и технологии. Область научного интереса – медицинская электроника и обработка медицинских сигналов

В.М. Бондарик

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

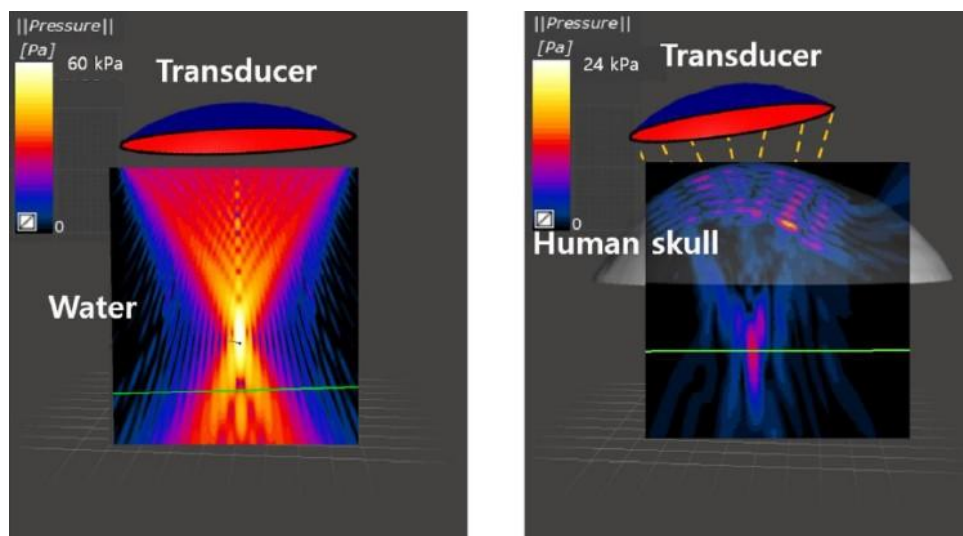
Аннотация. Рассмотрен пример обработки массива данных при процессе моделирования и симуляции в инфразвуковом диапазоне. Воздействие инфразвука на биологическую ткань.

Ключевые слова: *Big Data*, инфразвук, *SIMD*

Введение. Моделирование инфразвукового воздействия требует больших мощностей рабочей станции (PC), как правило, это многопроцессорные устройства, которые задействуют помимо центрального процессора, также и графический процессор [2]. Все это необходимо для построения модели биоткани и расчета конечного результата. Анализируя большие массивы данных, установив закономерности, можно создать модель объекта, которую можно использовать в самых разных целях [1]. С помощью модели можно изучать свойства уже описанного реального объекта, а также делать вероятностные прогнозы. В статье рассмотрен пример обработки массива данных при инфразвуковом воздействии на биоткань.

Основная часть. Моделирование инфразвукового воздействия производилось в программной среде *Sim4Life*. *Sim4Life* это среда моделирования и симуляции, совмещающая в себе множество модулей аналогичных биологическим тканям и воздействий на них.

Инфразвуковая волна, генерируемая композитным акустическим излучателем, имеет узкую полосу частот, достигающую нескольких Гц в одном импульсе [4]. Целью данного исследования было измерение характеристик передачи инфразвуковой волны, через фантом костной ткани и жидкой среды, и сравнение результатов с результатами численного моделирования. Средние плотности и скорости звука костной были рассчитаны на основе симуляции. Коэффициенты затухания варьировались в зависимости от места измерения, формы кости и толщины. На рисунке 1 видно, как инфразвуковая волна поглощается в воде (а), и в костной ткани (б).



а)

б)

Рисунок 1. а) поглощение инфразвука в воде, б) поглощение инфразвука в костной ткани.

Для того, чтобы провести моделирование воздействия инфразвука на биоткани, платформа *Sim4Life* обращается ко всем процессорным системам рабочей станции. Чаще всего это графические процессоры для многопоточковой обработки и анализа массива данных [3].

Процесс моделирования, заключающийся в потоковой обработке данных, известен под аббревиатурой *SIMD* (от англ. *Single Instruction – Multiple Data* – одна инструкция для множества данных, рисунок 2) [5]. Процессор, работающий по принципу *SIMD*, преобразует поток данных в поток результатов, используя программу как функцию преобразования

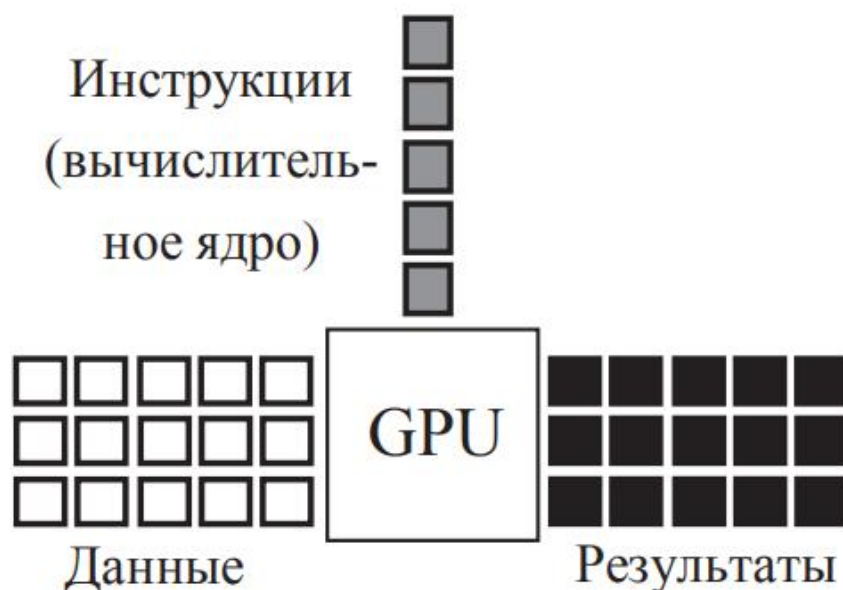


Рисунок 2. Поточно-параллельная обработка данных *SIMD*

Выбор концепции *SIMD* для графических процессоров обусловлен тем, что она обеспечивает параллельное использование большого количества «вычислителей» без явного управления ими: распределения задач, синхронизации вычислений и коммуникации между параллельными расчетами. Но графический процессор не имеет средств прямого взаимодействия с устройствами ввода-вывода (кроме монитора), а также доступа к оперативной памяти компьютера. Поэтому управление графическим процессором осуществляется только через центральный процессор [5]. Перед запуском программы, исполняемой на GPU, центральный процессор передает графическому процессору данные двух видов:

- значения констант, используемых в программе;
- один или несколько больших массивов данных для потоковой обработки.

Уже в первых графических процессорах была реализована конвейерная обработка потоков данных [5]. Заключается она в том, что каждый из параллельных «вычислителей» может одновременно обрабатывать несколько элементов потока, применяя к ним различные операции, например: загружать из памяти одно число, модифицировать другое и сохранять третье. Необходимо, чтобы операции, исполняемые конвейером одновременно, не зависели друг от друга, что как раз характерно для поточно-параллельных расчетов. Совершенствование вершинных и пиксельных шейдеров создало возможность программирования *GPU* для решения неграфических задач. Изначально при

этом вместо координат вершин требовалось передавать исходные числовые данные, алгоритм обработки которых программировали в форме вершинного шейдера. В настоящее время графические процессоры позволяют формулировать задачи общего назначения «естественным» для них образом, однако конвейерную обработку данных сохраняют. В качестве иллюстрации принципа конвейерной обработки данных на рисунке 3 показаны основные этапы графического конвейера, используемого графическим процессором визуализации трехмерных сцен.

Современные графические процессоры могут эффективно решать очень многие задачи. Основное требование к таким задачам – возможность распараллеливания по данным или по задачам, а также алгоритм, в котором вычисления превалируют над условными переходами и обменом данными между параллельными процессами. Наряду с симуляцией реальных процессов, GPU используются и для визуализации (отображение результатов томографии, представление сложных молекул и больших систем)

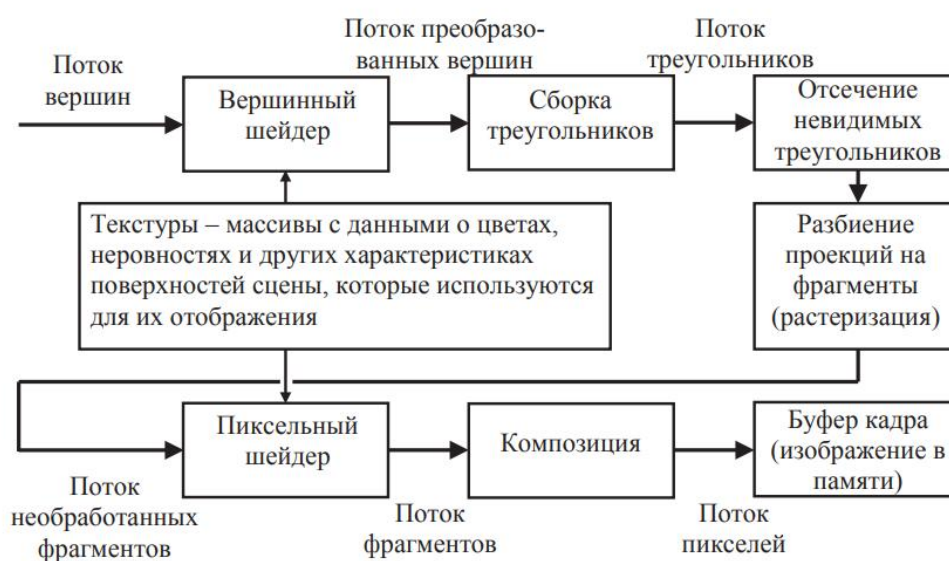


Рисунок 3. Этапы графического конвейера, выполняемые на графическом процессоре

Заключение. В статье были рассмотрены принципы обработки данных в процессе моделирования взаимодействия инфразвука с биологической тканью, в данном случае исследовалась костная ткань. Параллельная работа центрального и графического процессора, ускоряют обработку массива данных.

Список литературы

- [1] Трифонова О. П., Ильин В. А. «Большие данные» в биологии и медицине/Acta Naturae / Медицина и здравоохранение, ГРНТИ: Научная библиотека, 2013, № 3
- [2] Турлапов В., Гаврилов Н., Сапрыкин. В. Платформа цифровой медицины «Открытые системы», № 05, 2014 <http://www.osp.ru/os/2014/05/13041824>
- [3] Ильясова Н.Ю., Куприянов А.В., Попов С.Б., Парингер Р.А. «Особенности использования технологий Big Data в задачах медицинской диагностики. Системы высокой доступности» – 2016.
- [4] Сокол Г. И. «Особенности акустических процессов в инфразвуковом диапазоне частот». — Днепропетровск: Проминь, 2000. — 143 с.
- [5] Некрасов К.И., Поташников С.И., Боярченков А.С., Купряжкин А.Я., «Параллельные Вычисления Общего Назначения На Графических Процессорах» -2016.

Авторский вклад

Авторы внесли равный вклад в написание статьи.

**BIG DATA PROCESSING IN STUDYING THE INTERACTION OF
INFRASOUND WITH BIOLOGICAL TISSUE**

E.I. Leshchovich

*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*

Abstract. An example of processing a data array under infrasonic influence on biological tissue is considered.

Keywords: Big Data, infrasound, SIMD

УДК 338.1

БИЗНЕС-МОДЕЛИ СТАРТАПОВ: МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ВЫБОРА



И.В. Марахина

*Доцент кафедры экономики, БГУИР,
кандидат экономических наук, доцент
marahina@bsuir.by*

Марахина И.В.

Окончила Белорусский государственный университет информатики и радиоэлектроники. Более 15 лет работает в отраслях науки и образования. Сферы интересов: инновационное развитие, теория организации, маркетинг, поисковое продвижение и веб-аналитика.

Аннотация. Одним из важнейших решений в работе стартапа является выбор бизнес-модели. Сложность принятия такого решения определяется, с одной стороны, уникальностью и значимостью задачи, а с другой стороны – отсутствием в стартапе соответствующего специалиста для ее реализации. В представленной работе проанализированы существующие подходы к выбору и формированию бизнес-модели и выявлены недостатки их реализации в деятельности стартапов.

Ключевые слова: стартап, бизнес-модель, шаблон

Введение. В настоящее время одной из актуальных возможностей реализации направления коммерциализации научно-исследовательских работ в рамках модели «Университет 3.0» и обеспечения инновационного развития экономики в целом выступает активизация деятельности стартапов, которые являются точками роста инноваций, прорывных технологий и предпринимательской деятельности [1]. Следует отметить, что стартап имеет свою специфику, связанную с подходами к осуществлению деятельности, характеристиками продукта, высокими рисками и скоростями развития [2]. Можно говорить о высоком уровне уязвимости стартапов. Ранее уже были рассмотрены возможности минимизировать такие уязвимости [3]. В тоже время в статье основной акцент был сделан на развитие внешней среды.

Поэтому важным видится также исследование механизмов принятия важнейших управленческих решений в стартапе, связанных с выбором бизнес-модели. Такие решения по мнению ряда ученых в значительной степени влияют на успех фирмы. Например, [4] пишут, что по сути, многие известные истории успеха начинались с появления именно инновационной бизнес-модели, а не превосходного продукта. То есть часто оказывается, что в основе стартапа в первую очередь, даже не продукт, а новый подход к бизнесу.

Бизнес-модель рассматривается с двух сторон – компании и потребителя. В первом случае под бизнес-моделью, понимают способ получения эффекта от работы организации, стартапа. Во втором случае – способ создания ценности для потребителя. При этом чаще всего в определениях бизнес-модели оба эти направления совмещены.

– Бизнес-модель описывает обоснование того, как организация создает, доставляет и фиксирует ценность [5].

– Успешная инновационная бизнес-модель создает стоимость для клиентов и обеспечивает получение стоимости компанией [4, с. 23].

– Бизнес-модель стартапа представляет способ реализации стратегии формирования потребительской ценности, а также способ извлечения доходов от продаж [6, с. 55].

Следует отметить высокую важность разработок, направленных на помощь с выбором бизнес-модели, определяемую высокими темпами развития стартапа и необходимостью быстро принимать управленческие решения, с одной стороны, и нехваткой в команде специалистов в области экономики и менеджмента на ранних стадиях развития стартапа, с другой стороны.

Основная часть. В настоящее время представлено большое число трудов, содержащих различные бизнес-модели стартапа. Одной из популярных работ по исследуемой тематике является, книга О. Гассмана, К. Франкенбергера, М. Шика [4], где приводится 55 лучших по мнению авторов шаблонов бизнес-моделей: «Мы проанализировали наиболее революционные инновации в области бизнес-моделей за последние 50 лет и выявили предсказуемые и систематические шаблоны, лежащие в их основе. Наш метод построен на 55 шаблонах успешных бизнес-моделей, которые могут послужить вам образцом для собственных инновационных бизнес-моделей» [4]. Следует отметить, что указанные 55 шаблонов – это только вершина айсберга. Так, в работе указана необходимость их видоизменения и комбинирования в каждом случае. Кроме того, следует отметить, что существуют и другие бизнес-модели, которые авторы книги не отнесли к наиболее успешным.

Для формирования бизнес-модели в работе предлагают проанализировать четыре основных направления [4]: клиент, ценностное предложение, цепочка создания стоимости, механизм извлечения прибыли. Такой анализ включает ответ более чем на двадцать вопросов. Отметим, что одна компания может реализовывать сразу несколько шаблонов, о чем свидетельствуют в том числе и приводимые в книге примеры.

Подводя итоги, можно отметить, что представленные разработки затрагивают разные направления, разные аспекты деятельности организации и являются скорее набором эффективных практик, связанных с организацией производства, работой с поставщиками, потребителями. При этом при применении стартапами указанной методики выбора бизнес-модели можно выделить ряд недостатков:

– не все шаблоны подходят для стартапов – часть ориентирована на более зрелую стадию бизнеса;

– выбор модели требует перебора всех шаблонов и выработки ответов на основе приведенных в книге подсказок, что без соответствующих специалистов зачастую трудно осуществить;

– не выделены критерии классификации, позволяющие разделить, например, приводимые выше модели на группы, что позволило бы более эффективно с ними работать;

– бизнес-модели развиваются, и в работе ориентируют на поиск новой, интересной, прорывной бизнес-модели. То есть существует необходимость в дальнейшем развитии представленных в работе шаблонов моделей, что требует от стартапа значительных компетенций.

Также популярным является шаблон бизнес-модели предприятия (*Business Model Canvas*) А. Остервальдера и И. Пинье [5]. Он состоит из 9 блоков – ключевых элементов бизнеса. По каждому из блоков приводится перечень вариантов, которые может выбрать организация. При этом авторы указывают, что список приведенных вариантов (таблица 1) не является полным и исчерпывающим.

Таблица 1. Составляющие элементы шаблона бизнес-модели

<i>Составляющие</i>	<i>Варианты значений</i>
Потребительские сегменты (рынок)	массовый, нишевый, сегментация, разнообразный, многосторонний
Ценностные предложения	новизна, производительность, настройка, выполнение работы, бренд, статус, цена, снижение затрат, снижение рисков, доступность и удобство
Каналы привлечения потребителей	собственные, партнерские, смешанные
Организация взаимоотношений с потребителями	прямое взаимодействие, персонализированное обслуживание, самообслуживание, автоматизированное взаимоотношение, сообщества, совместное создание ценности
Поток доходов (за какую деятельность готов платить клиент)	продажа прав собственности на физический объект, плата за использование определенной услуги, абонентская плата, кредитование (аренда, лизинг), лицензирование, брокерские сборы, реклама
Ключевые виды ресурсов	физические, финансовые, интеллектуальная собственность и человеческие ресурсы
Ключевые виды деятельности	производство, персонализированное решение проблем клиентов, платформа (сеть).
Партнеры (разделяются по мотивации для создания взаимоотношений)	оптимизация и экономия за счет масштаба, снижение риска и неопределенности, приобретение определенных ресурсов и видов деятельности
Структура затрат	ориентированный на затраты и ценностно-ориентированные

Источник: [5].

Из минусов указанной бизнес-модели можно отметить:

- сложность реализации – необходимо провести значительную аналитическую работу для выработки бизнес-модели стартапа;
- сложность выбора наиболее правильного значения для каждого элемента шаблона;
- по мнению Б. Е. Токарева в модели «отсутствует важный элемент комплекса маркетинга – цена продукта» [6, с. 55-56].

В работе Э. Маурья представлен адаптированный «рациональный шаблон» (*lean canvas*) [7], дополненный такими элементами, как проблема и решение. Следует отметить, что и тут автор говорит о необходимости изменения бизнес-моделей [7]: «Я узнал напрямую, что, казалось бы, безупречные бизнес-модели могут распасться при выходе на новые рынки и под влиянием ожидания заинтересованных сторон». Э. Маурья большое значение уделяет таким характеристикам, как воспроизводство модели и обеспечение масштабирования. При этом он указывает на важность проверки экспоненциального роста ключевых характеристик на небольших временных промежутках (например, 3 месяца). Таким образом, при таком подходе представленный шаблон позволит сформировать бизнес-модель стартапа, но требует квалифицированной помощи на протяжении всей работы стартапа.

Заключение. В работе были проанализированы подходы к формированию бизнес-модели стартапа и выделены их недостатки. В первую очередь они определяются сложностью выбора модели, необходимостью анализа большого объема информации и привлечения к выбору экспертов, которые могли бы проработать метрики и сделать наиболее эффективный выбор. В тоже время зачастую у стартапа доступ к таким экспертам ограничен. Таким образом, важным видится формирование алгоритма, который позволил бы стартапам без привлечения экспертов выбирать наиболее перспективные бизнес-модели.

Список литературы

- [1] Марахина, И. В. Инвестиционная инфраструктура белорусской стартап-экосистемы: субъекты, проблемы и направления развития / И. В. Марахина // Банковский вестник. – 2020. – №7. – С. 49–58
- [2] Марахина, И. В. Этапы жизненного цикла стартапа: сущность, признаки, проблемы и пути их предотвращения / И. В. Марахина // Цифровая трансформация. – 2021. – № 2 (15). – С. 21–31
- [3] Марахина, И. В. Теоретические аспекты функционирования стартап-экосистемы/ И. В. Марахина // Цифровая трансформация. – 2020. – № 3 (12). – С. 31–38.
- [4] Гасман, О. Бизнес-модели: 55 лучших шаблонов / О. Гасман, К. Франкенбергер, М. Шик. – «Альпина Диджитал», 2014.
- [5] Osterwalder, A. Business Model Generation / A. Osterwalder, Y. Pigneur. – Amsterdam, The Netherlands, 2009. – 249 p.
- [6] Токарев, Б. Е. Модель бизнеса стартапа инновационного продукта / Б. Е. Токарев // Вестник университета № 9, 2018. – 2018. – № 9. – С. 55–58.
- [7] Maurya, A. Scaling Lean: Mastering the Key Metrics for Startup Growth / A. Maurya. – Portfolio/Penguin, 2016. – 290 p.

BUSINESS MODELS OF STARTUPS: METHODOLOGICAL BASIS OF SELECTION

I.V. Marahina,

*Associate professor of the Department of Economics of BSUIR,
Candidate of Economic Sciences, Associate Professor*

Annotation. One of the most important decisions in a startup is the choice of business model. The difficulty of making such a decision is determined, on the one hand, by the uniqueness and significance of the task, and on the other hand, by the lack of an appropriate specialist in the startup to implement it. The presented work analyzes existing approaches to the selection and formation of a business model and identifies the shortcomings of their use in the activities of startups.

Keywords: startup, business model, template

УДК 004.032:004.75

АНСАМБЛИ НЕЙРОННЫХ СЕТЕЙ ДЛЯ КЛАССИФИКАЦИИ СОСТОЯНИЙ И РЕЖИМОВ ФУНКЦИОНИРОВАНИЯ ПОДСИСТЕМ КОСМИЧЕСКОГО АППАРАТА



М.М. Лукашевич
Доцент кафедры информационных систем управления БГУ, кандидат технических наук, доцент
LukashevichMM@bsu.by



В.Ю. Скобцов
Доцент кафедры компьютерных технологий и программной инженерии Санкт-Петербургского государственного университета аэрокосмического приборостроения, кандидат технических наук, доцент
vasko_vasko@mail.ru
и



А.В. Инютин
Зав. лабораторией №222 ОИПИ,
avin@lsi.bas-net.by



В.В. Ганченко
ст. науч. сотр. лаборатории №222 ОИПИ, кандидат технических наук,
ganchenko@lsi.bas-net.by

М.М. Лукашевич

Окончила учреждение образования «Белорусский государственный университет информатики и радиоэлектроники». Область научных интересов связана с разработкой методов и алгоритмов машинного обучения для решения задач анализа данных и компьютерного зрения.

В.Ю. Скобцов

Окончил Донецкий государственный университет. Область научных интересов связана с разработкой и исследованием методов интеллектуального анализа данных, машинного обучения, мягких вычислений, логико-вероятностного моделирования и их применением в области анализа данных функционирования информационных и аппаратных систем, их надежности и живучести, диагностики программно-аппаратных систем.

А.В. Инютин

Окончил Белорусский государственный университет транспорта. Заведующий лабораторией идентификации систем ОИПИ НАН Беларуси. Проводит научные исследования в области обработки и распознавания изображений методами математической морфологии, обработки данных ДЗЗ и телеметрии космических аппаратов.

В.В. Ганченко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с применением компьютерного зрения и машинного обучения в задачах обработки изображений.

Аннотация. Выполнен разведочный анализ и подготовка наборов данных для решения задач классификации состояний подсистемы КА. Рассмотрены и реализованы ансамблевые методы машинного обучения для решения задач классификации состояния подсистем космического аппарата (КА). Полученные результаты носят прикладной характер и являются элементом комплексной автоматизации и интеллектуализации процессов оценивания и анализа данных телеметрической информации МКА (малого космического аппарата).

Ключевые слова: нейронные сети, ансамблевые методы, классификация, бортовая аппаратура (БА), телеметрическая информация (ТМИ), малый космический аппарат (МКА).

Введение. Одной из важнейших задач на всех этапах жизненного цикла малого космического аппарата (МКА) является анализ их телеметрической информации о функционировании бортовой аппаратуры (БА) МКА с точки зрения определения их технического состояния для обеспечения их штатного и безопасного функционирования. Актуальность обусловлена прежде всего тем, что одной из основных причин потерь МКА являются отказы, сбои и некорректная работа систем БА МКА. Большое количество информации, поступающей и накапливающейся в специализированных банках данных о БА МКА, может быть эффективно использовано для совершенствования процесса определения технического состояния МКА. В связи с тем, что данные о функционировании БА МКА, включающие телеметрические измерения, представляют собой разнородные нерегулярные многомерные данные, актуальным является исследование, разработка и применение моделей, которые позволяют анализировать такого рода данные с возможностью извлечения из них полезной информации и последующего построения с их использованием классификационных и прогностических моделей с целью определения технического состояния МКА для принятия корректных управляющих и эксплуатационных решений в процессе функционирования МКА. Методы машинного обучения, искусственного интеллекта и математической статистики являются на текущий момент одними из самых перспективных и широко используемых подходов в анализе данных высокотехнологичных систем. В современных условиях для решения перечисленных задач, обеспечения требуемой степени автономности, качества и оперативности управления такими сложными объектами как МКА необходимо выполнить комплексную автоматизацию и интеллектуализацию процессов оценивания и многомодельного анализа данных телеметрической информации МКА. Однако, в большинстве случаев на практике автоматизация выполнена, в лучшем случае, лишь частично, и многое делается зачастую вручную, базируясь на эвристических правилах [1]. Таким образом, задача интеллектуального анализа данных телеметрии БА МКА с целью определения технического состояния БА МКА является актуальной и востребованной [1, 2].

Целью исследования являлась разработка моделей машинного обучения для классификации состояния и режимов функционирования подсистем космического аппарата (КА).

Разведочный анализ данных. В качестве данных для построения моделей машинного обучения были использованы данные телеметрической информации (ТМИ) звездного датчика (ЗД) БА Белорусского космического аппарата (БКА) за 2015, 2021 годы. Перечень признаков ТМИ ЗД БКА приведен в таблице 1.

Таблица 1. Перечень телеметрической информации ЗД БКА

Название датчика	Значение датчика
УАС	координаты по солнцу
УВС	координаты по солнцу
ТОК58	ток на ЗД
STRQ1-1	значения кватернионов
STRQ2-1	значения кватернионов
STRQ3-1	значения кватернионов
STRQ4-1	значения кватернионов
TSTR1DSP	температура процессора обработки сигнала
TSTR1CCD	температура ПЗС матрицы (которая делает снимки)
Class1	0 – нештатного состояния подсистемы 1 – штатного состояния подсистемы
Class	0 – завершено 2 – инициализация разворота 3 – ожидание разворота 4 – разворот 8 – отмена

Данные сохранены в файл в формате .csv, всего 2 384 627 объектов, описанных 11 признаками (столбцами). Два столбца («Class1», «Class») являются целевыми или метками классов для 2-х классовых задач и для 5-ти классовых задач классификации состояния подсистемы. Для набора данных выполнен разведочный анализ данных (EDA, *exploration data analysis*) с использованием библиотеки pandas. Результаты вывода первых 5 строк представлены на рисунке 1.

	УВС	УАС	ТОК58	STRQ1	STRQ2	STRQ3	STRQ4	TSTR1CCD	TSTR1DSP	Class	Class1
0	0.000	0.000	139.0	0.137	-0.312	0.78	0.524	9.849	23.41	0	0
1	0.000	0.000	135.0	0.137	-0.312	0.78	0.524	9.849	23.41	0	0
2	0.000	0.000	135.0	0.132	-0.315	0.78	0.525	9.849	23.41	0	0
3	46.865	32.324	135.0	0.132	-0.315	0.78	0.525	9.849	23.41	0	0
4	21.464	19.704	135.0	0.132	-0.315	0.78	0.525	9.849	23.41	0	0

Рисунок 1. Вывод первых 5 строк таблицы данных

Общая информация о наборе данных, тип каждого признака, а также наличие или отсутствие пропусков представлены на рисунке 2. Статистическая информация по каждому числовому признаку представлена на рисунке 3. Визуализация матрицы корреляции признаков представлена на рисунке 4.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2384627 entries, 0 to 2384626
Data columns (total 11 columns):
#   Column      Dtype
---  ---
0   YBC         float64
1   YAC         float64
2   TOK58       float64
3   STRQ1       float64
4   STRQ2       float64
5   STRQ3       float64
6   STRQ4       float64
7   TSTR1CCD    float64
8   TSTR1DSP    float64
9   Class       int64
10  Class1      int64
dtypes: float64(9), int64(2)
memory usage: 200.1 MB
```

Рисунок 2. Общая информация о наборе данных, тип каждого признака, а также наличие или отсутствие пропусков

	YBC	YAC	TOK58	STRQ1	STRQ2	STRQ3	STRQ4	TSTR1CCD	TSTR1DSP	Class	Class1
count	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06	2.384627e+06
mean	-1.631350e+01	1.201425e+01	8.774924e+01	2.154371e-01	2.675968e-02	3.409743e-01	4.153858e-01	1.090822e+01	2.866912e+01	7.521285e-01	1.138807e-01
std	4.555501e+01	1.849339e+01	1.968790e+02	3.241060e-01	4.108809e-01	4.478304e-01	4.317669e-01	3.391021e+03	2.563074e+03	2.146762e+00	3.176664e-01
min	-9.000000e+00	-9.000000e+01	0.000000e+00	-7.050000e-01	-7.060000e-01	-7.030000e-01	-7.040000e-01	-4.121000e+01	-4.121000e+01	0.000000e+00	0.000000e+00
25%	-5.247600e+01	3.594000e+00	0.000000e+00	9.000000e-03	-3.160000e-01	-5.000000e-03	1.310000e-01	3.064000e+00	2.341000e+01	0.000000e+00	0.000000e+00
50%	-2.725700e+01	1.100800e+01	1.330000e+02	2.550000e-01	-9.100000e-02	4.410000e-01	5.370000e-01	6.969000e+00	2.471400e+01	0.000000e+00	0.000000e+00
75%	1.650900e+01	1.970000e+01	1.380000e+02	3.990000e-01	3.250000e-01	7.330000e-01	7.810000e-01	1.102200e+01	2.669400e+01	0.000000e+00	0.000000e+00
max	9.000000e+01	9.000000e+01	6.543300e+04	9.990000e-01	1.000000e+00	1.000000e+00	1.000000e+00	3.177054e+06	2.296978e+06	8.000000e+00	1.000000e+00

Рисунок 3. Статистическая информация по каждому числовому признаку

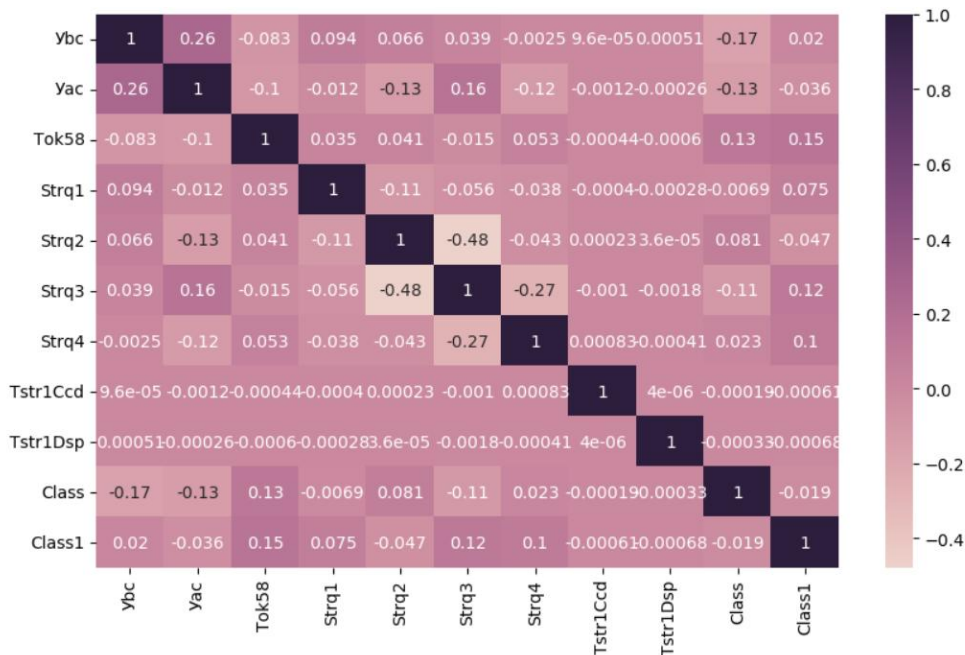


Рисунок 4. Визуализация матрицы корреляции признаков

Описание целевых признаков: а) уникальные значения целевых признаков, число объектов с уникальными значениями для целевых признаков представлено на рисунке 5.

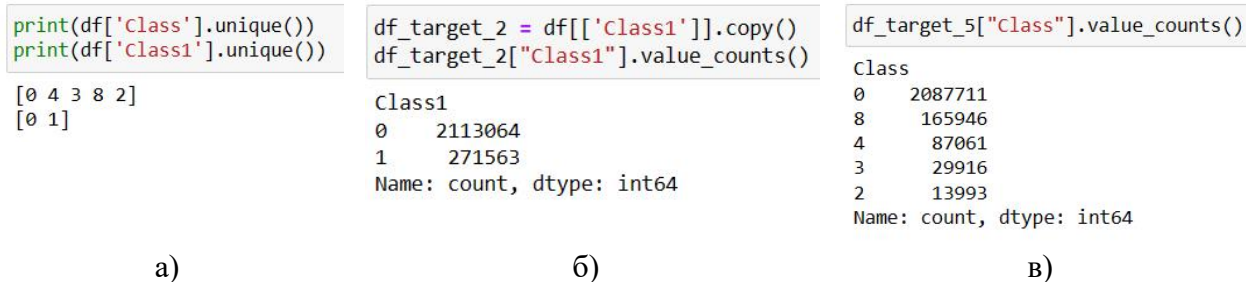


Рисунок 5. Описание целевых признаков: а) уникальные значения целевых признаков, б) и в) число объектов с уникальными значениями для целевых признаков

В результате разведочного анализа данных установлено, что пропущенные значения признаков отсутствуют, все признаки являются числовыми, набор данных с учётом меток классов («Class», «Class1») является несбалансированными. Метки «Class» для дальнейшего этапа построения классификаторов требуют перекодировки значений, таблица 2.

Таблица 2. Перекодировка меток для «Class»

«Class»	Режим подсистемы ориентации и стабилизации	Перекодировка классов	Число объектов
0	Завершено	0	2 087 711
8	Отмена	1	165 946
4	Разворот	2	87 061
3	Ожидание разворота	3	29 916
2	Инициализация разворота	4	13 993

Балансировка данных. В задаче классификации данные называются несбалансированными (*Imbalanced Data*), если в обучающей выборке доли объектов разных классов существенно различаются, также говорят, что классы не сбалансированы. Для решения существующей проблемы было использовано два подхода: замена большого класса подвыборкой по мощности равной малому классу (недосэмплированием, *Undersampling the majority class*) и увеличение в размерах малого класса (пересемплирование, *Oversampling the minority class*). Простейшая стратегия недосэмплирования – взять случайную подвыборку, простейшая стратегия пересэмплирования – продублировать объекты малого класса [3-5], рисунок 10.

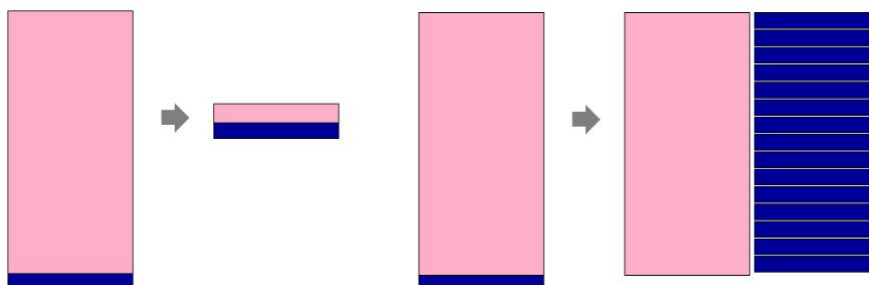


Рисунок 6. *Undersampling* (слева) и *Oversampling* (справа)

Подготовка данных для обучения ансамблей нейронных сетей включала в себя подготовку двух набора для 2-классовой задачи классификации и 5-ти классовой задача классификации с назначением столбцов «Class1» и «Class» целевыми столбцами соответственно («Набор данных 1» и «Набор данных 2»). Далее выполнялось разбиение

наборов данных на обучающую и тестовую выборки в соотношении 70% и 30% соответственно. Балансировка данных выполнялась только для обучающих наборов, чтобы получить независимые оценки точности классификации на тестовых наборах и избежать «утечки данных». Недосэмплирование и пересемплирование выполнялось средствами библиотеки *sklearn*. В случае «Набора данных 1» выполнялось недосемплирование для класса «0» и пересемплирование для класса «1». Статистика по исходному «Набору данных 1» представлена в таблице 3, результаты разбиение на обучающую и тестовую выборки - в таблице 4, результаты балансировки – в таблице 5.

Таблица 3. Исходных «Набор данных 1» (2 класса)

Класс	Число объектов
0	2 113 064
1	271 563

Таблица 4. Результаты разбиение на обучающую и тестовую выборки «Набора данных 1»

Набор данных, число объектов	Обучающая выборка, число объектов	Тестовая выборка, число объектов
2 384 627	1 669 238	715 389
0 класс – 2 113 064 1 класс – 271 563	0 класс – 1 479 190 1 класс – 190 048	0 класс – 633 874 1 класс - 81 515

Таблица 5. Результаты балансировки обучающей выборки «Набора данных 1»

Класс	Число объектов после балансировки	Использованная техника
0 класс	190 094	уменьшено число образцов
1 класс	190 094	увеличено число образцов

Для «Набора данных 2» выполнялось недосемплирование для класса «0» и пересемплирование для классов «1» и «2». Статистика по исходному «Набору данных 2» представлена в таблице 6, результаты разбиение на обучающую и тестовую выборки – в таблице 7, результаты балансировки – в таблице 8.

Таблица 6. Исходных «Набор данных 2» (5 класса)

Класс	Число объектов
0	1 461 397
1	9 795
2	20 940
3	60 943
4	116 162

Таблица 7. Результаты разбиение на обучающую и тестовую выборки «Набора данных 2»

Набор данных, число объектов	Обучающая выборка, число объектов	Тестовая выборка, число объектов
2 384 627	1 669 238	715 389
0 класс – 2 087 711 1 класс – 13 993 2 класс – 29 916 3 класс – 87 061 4 класс – 165 946	0 класс – 1 461 397 1 класс – 9 795 2 класс – 20 941 3 класс – 60 943 4 класс – 116 162	0 класс – 626 314 1 класс – 4 198 2 класс – 8 975 3 класс – 26 118 4 класс – 49 784

Таблица 8. Результаты балансировки обучающей выборки «Набора данных 2»

Класс	Число объектов после балансировки	Использованная техника
0 класс	116 162	уменьшено число образцов
1 класс	19 590	увеличено число образцов
2 класс	60 943	увеличено число образцов
3 класс	60 943	без изменений
4 класс	116 162	без изменений

Ансамблевые методы машинного обучения. Ансамблевые методы являются мощным инструментом машинного обучения и находят широкое применение. Цель ансамблевых методов – объединить прогнозы нескольких базовых оценок, построенных с заданным алгоритмом обучения, чтобы улучшить обобщаемость / надежность по сравнению с одной оценкой. Модели нулевого уровня (*базовые модели*) – это модели, обученные на тренировочных данных, прогнозы которых собираются. Модель первого уровня (*метамодель*) – это модель, которая учится наилучшим образом комбинировать прогнозы базовых моделей.

В методах усреднения главный принцип состоит в том, чтобы построить несколько оценщиков независимо, а затем усреднить их прогнозы. Комбинированная оценка обычно лучше, чем любая из оценок с одной базой, потому что ее дисперсия уменьшается. В методах усиления базовые оценки строятся последовательно, и каждый пытается уменьшить смещение комбинированной оценки. Мотивация состоит в том, чтобы объединить несколько слабых моделей для создания мощного ансамбля. Наиболее широко используются следующие ансамблевые методы: бэггинг (модельное усреднение), бустинг (модельное усиление) и стэкинг (модельное наложение) [6-9]. В рамках реализации этапа проекта было рассмотрено два метода построения ансамблей моделей: бэггинг (модельное усреднение) и мягкое и жесткое голосование.

Бэггинг (*bagging, bootstrap aggregation*), или модельное усреднение, – это разновидность параллельного ансамблирования, которая используется для урегулирования проблем, связанных с высокой дисперсией в моделях машинного обучения. Бутстраповская часть бэггинга относится к наборам данных, используемым для тренировки членов ансамбля. В частности, если существует k подмоделей, то для тренировки каждой подмодели ансамбля используется k отдельных наборов данных. Каждый набор данных строится путем случайного отбора экземпляров из изначального тренировочного набора данных (с возвратом). Это означает, что существует высокая вероятность того, что в любом из k наборов данных будут отсутствовать некоторые тренировочные примеры, но и любой набор данных, вероятно, будет иметь повторяющиеся тренировочные примеры. Агрегирование выполняется на результате работы многочисленных членов ансамблевой модели – берется либо среднее значение в случае регрессионной задачи, либо большинство голосов в случае классификационной [10-14].



Рисунок 7. Схематичное представление бэггинга

Жесткое голосование и мягкое голосование – это два разных способа объединения предсказаний нескольких классификаторов в ансамблевых алгоритмах машинного обучения. При жестком голосовании каждый классификатор делает свое предсказание, а предсказание ансамбля – это просто большинство голосов. Мягкое голосование более точное, но и более сложное в реализации. Выбор оптимального подхода зависит от конкретного приложения.

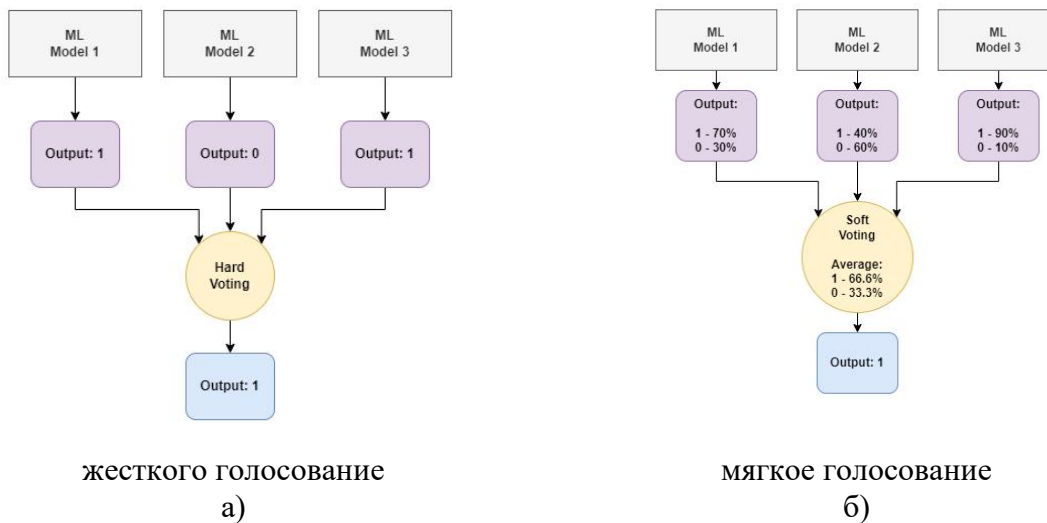


Рисунок 8. Схематичное представление а) – жесткого и б) – мягкого голосования

Рассмотренные методы были использованы при реализации ансамблей нейронных сетей. Искусственные нейронные сети являются современным подходом машинного обучения. Нейронные сети находят эффективное применение в задачах анализа различных типов данных: изображений, видеорядов, текстовых и аудио данных, в том числе в задачах классификации данных временных рядов различной природы [15]. Базовая модель нейронной сети была предложена на предыдущих этапах исследования и использована при построении ансамблей нейронных сетей. Базовая модель нейронной сети является гибридной классификационной моделью, включающей свёрточные и полносвязные слои, слои субдекретизации также блоки *GRU (Gated Recurrent Units)*, использованы техники регуляризации [16-18]. Обозначим базовую модель как «*CNN+GRU+Dense* архитектура».

Разработка ансамблей нейронных сетей для классификации режимов функционирования подсистем КА. В случае решения задачи 2х классовой классификации выход нейронной сети содержит 1 нейрон и используется сигмоидная функция активации.

Здесь и далее в качестве оптимизатора был использован алгоритм *Adam*. Результаты этого оптимизатора, как правило, лучше, чем у любого другого алгоритма оптимизации, занимают меньшее время и требуют меньше параметров для настройки. Благодаря этому метод *Adam* рекомендуется в качестве оптимизатора по умолчанию для большинства приложений обучения современных моделей нейронных сетей. Общее число тренировочных объектов, представленных в одном мини-батче равно 1 024, число эпох обучения 100. В процессе обучения использовались механизмы мониторинга модели с использованием метрики точности, а также механизм ранней остановки с целью исключения переобучения модели.

Для реализации нейронных сетей и механизмов их обучения использовались язык *Python*, фреймворк глубокого обучения *Keras*, а также библиотека машинного обучения *sklearn*. При реализации ансамблей нейронных сетей использовалась библиотека *SciKeras*.

Для оценки эффективности применения ансамблевых методов в начале была реализована и обучена базовая модель, а затем эта модель использовалась при реализации ансамблевых моделей методом бэггинга и методом жёсткого голосования. Во всех экспериментах число базовых моделей в ансамбле равно 3. Результаты тестирования (классификационные отчеты) моделей на тестовом наборе из «Набора данных 1» представлены в таблице 9.

Таблица 9. Результаты тестирования (классификационные отчеты) моделей на тестовом наборе из «Набора данных 1»

Архитектура/ансамбль	Классификационный отчет для тестового набора				
	precision	recall	f1-score	support	
Одна « <i>CNN+GRU+Dense</i> архитектура»	0	0.99	0.82	0.90	633920
	1	0.40	0.96	0.56	81469
	accuracy			0.83	715389
	macro avg	0.70	0.89	0.73	715389
	weighted avg	0.93	0.83	0.86	715389
<i>Voting «hard»</i> три « <i>CNN+GRU+Dense</i> архитектуры»	0	0.99	0.82	0.90	633920
	1	0.41	0.97	0.58	81469
	accuracy			0.84	715389
	macro avg	0.70	0.89	0.74	715389
	weighted avg	0.93	0.84	0.86	715389
Bagging три « <i>CNN+GRU+Dense</i> архитектуры»	0	0.99	0.77	0.87	633920
	1	0.35	0.97	0.52	81469
	accuracy			0.79	715389
	macro avg	0.67	0.87	0.69	715389
	weighted avg	0.92	0.79	0.83	715389

В результате экспериментов установлено, что ансамбль, реализованный на основе метода жёсткого голосования позволяет увеличить точность (accuracy) классификации на 1% в сравнение с базовой моделью, с 83% до 84% соответственно для двухклассовой задачи классификации.

В случае решения задачи 5ти классовой классификации выход нейронной сети содержит 5 нейронов и используется функция активации *softmax*. Для оценки

эффективности применения ансамблевых методов в начале была реализована и обучена базовая модель, а затем эта модель использовалась при реализации ансамблевых моделей методом бэггинга и методом мягкого голосования. Во всех экспериментах число базовых моделей в ансамбле равно 3. Результаты тестирования (классификационные отчеты) моделей на тестовом наборе из «Набора данных 2» представлены в таблице 10.

Таблица 10. Результаты тестирования (классификационные отчеты) моделей на тестовом наборе из «Набора данных 2»

Архитектура/ансамбль	Классификационный отчет для тестового набора				
Одна «CNN+GRU+Dense архитектура»		precision	recall	f1-score	support
	0	0.99	0.79	0.88	626314
	1	0.62	0.70	0.66	4198
	2	0.19	0.58	0.28	8975
	3	0.18	0.60	0.28	26118
	4	0.47	0.93	0.63	49784
	accuracy			0.79	715389
	macro avg	0.49	0.72	0.54	715389
	weighted avg	0.91	0.79	0.83	715389
	Voting «soft» три «CNN+GRU+Dense архитектуры»		precision	recall	f1-score
0		0.99	0.79	0.88	626314
1		0.61	0.72	0.66	4198
2		0.20	0.68	0.31	8975
3		0.22	0.58	0.32	26118
4		0.45	0.97	0.61	49784
accuracy				0.80	715389
macro avg		0.49	0.75	0.56	715389
weighted avg		0.91	0.80	0.83	715389
Bagging три «CNN+GRU+Dense архитектуры»			precision	recall	f1-score
	0	0.99	0.78	0.87	626314
	1	0.62	0.69	0.66	4198
	2	0.19	0.62	0.29	8975
	3	0.19	0.58	0.29	26118
	4	0.43	0.96	0.60	49784
	accuracy			0.78	715389
	macro avg	0.48	0.72	0.54	715389
	weighted avg	0.91	0.78	0.82	715389

В результате экспериментов установлено, что ансамбль, реализованный на основе метода мягкого голосования позволяет увеличить точность (*accuracy*) классификации на 1% в сравнение с базовой моделью, с 79% до 80% соответственно для пятиклассовой задачи классификации.

Заключение. В результате исследований выполнен разведочный анализ и подготовка наборов данных для решения задач классификации состояний подсистемы КА. Рассмотрены и реализованы ансамблевые методы машинного обучения (бэггинг и голосование) для решения задач классификации состояния подсистем КА (2 класса и 5 классов). Установлено, что ансамбль нейронных сетей, реализованный на основе метода жёсткого голосования позволяет увеличить точность (*accuracy*) классификации (2 класса) состояния подсистемы КА 1% в сравнение с базовой моделью, с 83% до 84% соответственно. Определено, что ансамбль, реализованный на основе метода мягкого голосования, позволяет увеличить точность (*accuracy*) классификации (5 классов) режимов функционирования подсистемы КА на 1% в сравнение с базовой моделью, с 79% до 80% соответственно. Полученные результаты носят прикладной характер и являются элементом комплексной автоматизации и интеллектуализацию процессов оценивания и многомодельного анализа данных телеметрической информации МКА.

Дальнейшие исследования целесообразно продолжить в части модификации использованной и поиска новых эффективных архитектур нейронных сетей для

построения ансамблей и инжиниринга новых признаков, описывающих состояние подсистемы КА в виду их ограниченного числа.

Список литературы

- [1] Концепция проактивного управления сложными объектами: теоретические и технологические основы / М. Ю. Охтилев [и др.] // Известия вузов. Приборостроение. – 2014. – Т. 57, №11. – С.7–14.
- [2] Проект стратегии информационных технологий Госкорпорации «Роскосмос». – [Электронный документ]. – Режим доступа: <https://www.roscosmos.ru/25892/> (дата обращения: 07.06.2023).
- [3] CSMOUTE: Combined Synthetic Oversampling and Undersampling Technique for Imbalanced Data Classification [Электронный документ]. -- Режим доступа: <https://arxiv.org/pdf/2004.03409.pdf> (дата обращения: 02.11.2023).
- [4] Deep Synthetic Minority Over-Sampling Technique [Электронный документ]. -- Режим доступа: <https://arxiv.org/ftp/arxiv/papers/2003/2003.09788.pdf> (дата обращения: 04.12.2023).
- [5] Automatic Feature Selection [Электронный документ]. – Режим доступа: // <https://amueller.github.io/aml/05-advanced-topics/12-feature-selection.html> (дата обращения: 04.12.2023).
- [6] Брейман Л. Вставка небольших голосов для классификации в больших базах данных и в Интернете / Л. Брейман // Машинное обучение. – 1999. – Том 36, № 1. – С. 85-103.
- [7] Брейман Л. Предикторы упаковки / Л. Брейман // Машинное обучение. – 1996. – Том 24, № 2. – С. 123-140.
- [8] Хо Т. Метод случайных подпространств для построения лесов решений / Т. Хо // Анализ шаблонов и машинный интеллект. – 1998. – Том 20, № 8. – С. 832-844.
- [9] Луппе Г. Ансамбли на случайных участках / Г. Луппе и П. Геуртс // Машинное обучение и обнаружение знаний в базах данных. – 2012. – С. 346-361.
- [10] Tibshirani R. The Elements of Statistical Learning / Tibshirani, R., Friedman, J. – 2001. – Springer, New York. – 764 p.
- [11] Domingos P. A Unified Bias-Variance Decomposition and its Applications / P. Domingos // Proc. 17th International Conf. on Machine Learning. – 2000. – PP. 231-238.
- [12] Breiman L. Random Forests / Breiman L. // Machine Learning. – 2001. – Vol. 451. PP. 5–32.
- [13] Власенко М. Ансамблевые методы машинного обучения в скоринге розничных клиентов [Электронный документ]. -- Режим доступа: <https://www.nbrb.by/bv/pdf/articles/11056.pdf> (дата обращения: 20.10.2023).
- [14] Вольперт Д.Х. Сложное обобщение / Дэвид Х.Вольперт // Нейронные сети. – 1992. – Том 5, №2. – С. 241-259.
- [15] Fawaz H. I. et al. (2019) Muller Deep learning for time series classification: a review. Data Mining and Knowledge Discovery. 33, 917–963.
- [16] Skobtsov V.Y. Automatic Searching the Neural Network Models for Time Series Classification of Small Spacecraft's Telemetry Data with Genetic Algorithms / Skobtsov, V.Y., Stasiuk, A. // Artificial Intelligence Application in Networks and Systems. CSOC 2023. Lecture Notes in Networks and Systems. – 2023. – Vol. 724. – PP. 800–811. doi.org/10.1007/978-3-031-35314-7_66.
- [17] Skobtsov V.Yu. Binary classification of small satellites telemetry data based on deep learning approach // Applied Aspects of Information Technology. 2021. Vol. 4 No.4, pp. 299–310.
- [18] В.Ю.Скобцов, Б.В.Соколов Гибридные нейросетевые модели в задаче мультиклассовой классификации данных телеметрической информации малых космических аппаратов // Вестник ВГУ. Системный анализ и информационные технологии, 2022, № 3, С.99-114

Авторский вклад

Лукашевич Марина Михайловна – разработка нейросетевых классификаторов на основе ансамблевых методов, подготовка и балансировка наборов данных.

Скобцов Вадим Юрьевич – разработка базовой нейросетевой модели, сравнение результатов классификации, анализ полученных результатов.

Инютин Александр Владимирович – руководитель проекта, постановка задачи исследования.

Ганченко Валентин Вячеславович – разведочный анализ данных.

NEURAL NETWORKS ENSEMBLES FOR CLASSIFICATION OF OPERATION STATES AND MODES OF SPACECRAFT SUBSYSTEMS

M.M. Lukashevich
*Associate Professor,
Department of
Information Systems
Management, BSU, PhD
of Technical sciences,
Associate Professor*

V.Yu. Skobtsov
*Associate Professor,
Department of
Computer Technology
and Software
Engineering, Saint
Petersburg State
University of Aerospace
Instrumentation, PhD,
Associate Professor*

A.V. Inyutin
*Head of Laboratory,
UIIP*

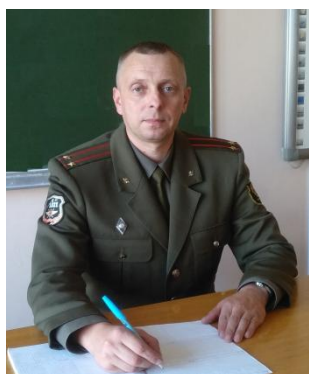
V.V. Ganchenko
*Researcher, UIIP, PhD
of Technical sciences*

Abstract. Exploratory analysis and preparation of datasets for solving spacecraft subsystem state and modes classification problems are performed. Ensemble methods of machine learning for solving the tasks of classifying the state and mode of spacecraft subsystems are considered and realized. The obtained results are of applied nature and are an element of complex automation and intellectualization of the processes of evaluation and analysis of small spacecraft telemetry data.

Keywords: neural networks, ensemble methods, classification, on-board equipment, telemetry information, small spacecraft.

УДК 004.62

BIG DATA В АВТОМОБИЛЬНОЙ ПРОМЫШЛЕННОСТИ



А.А.Сасновский
Магистр,
старший преподаватель
кафедры связи военного
факультета БГУИР
a.sasnovski@bsuir.by



Р.А.Божко
Магистр,
начальник цикла
кафедры связи военного
факультета БГУИР
r.bozhko@bsuir.by

А.А.Сасновский

Окончил «Военную академию Республики Беларусь». Магистр. Работает в БГУИР в должности старшего преподавателя.

Р.А.Божко

Окончил «Военную академию Республики Беларусь». Магистр. Работает в БГУИР в должности начальника цикла кафедры связи военного факультета.

Аннотация. В этой статье рассмотрена значимость использования больших данных в различных сферах.

Ключевые слова: большие данные, информационные технологии, интернет-вещей.

Введение. Компьютеры, всемирная сеть и современные информационные технологии оказывают огромное влияние на развитие современного общества. Огромные потоки разнообразной, а порой и противоречивой информации, неоспоримо оказывают влияние на социально-политические и экономические процессы. Только тот, кто владеет информацией, способен оперативно оценить обстановку, проанализировать развитие ситуации и принять правильное решение имеет огромное преимущество перед другими. Для получения информации пользователь при помощи специальных автоматизированных инструментов находит, а порой и хранит, в огромных хранилищах данных интересующую его информацию чтобы использовать для статистики, анализа, прогнозов и принятия решений. С каждым днем массивы данных постоянно увеличиваются, а процесс поиска вызывает все больше затруднений ввиду разнообразия хранящихся данных. Они в своем составе содержат как обучающую информацию, так и данные о работе разнообразных устройств управлять которыми можно удаленно.

Современные вычислительные системы обеспечивают мгновенный доступ к массивам больших данных. Для их хранения используют специальные дата-центры с самыми мощными серверами. Благодаря высокопроизводительным технологиям компании могут использовать любые объемы больших данных для анализа отбирая только те, что нужны для анализа. Все чаще большие данные применяют для задач в рамках расширенной аналитики используя методы описательной, прогнозной

предписательной и диагностической аналитики [1].

Анализ применения больших данных. Анализ больших данных широко развивается в государственном управлении, промышленности, медицине и системе интернет вещей и других отраслях. Изучение и анализ больших данных в государственном управлении помогает принимать решения в таких областях, как здравоохранение, занятость населения, экономическое регулирование, борьба с преступностью и обеспечение безопасности, реагирование на чрезвычайные ситуации. В промышленности анализ больших данных повышает прозрачность промышленных процессов и позволяет более точно планировать расходование ресурсов. В медицине большие данные помогают точнее ставить диагнозы и подбирать эффективное лечение.

Огромное развитие в настоящее время уделяется «системе» интернет вещей. Большие данные и интернет вещей неразрывно связаны между собой. Промышленные и бытовые приборы, подключенные к интернету вещей, собирают огромное количество данных, на основе анализа которых впоследствии регулируется работа этих приборов [2].

Интернет вещей – это система взаимосвязанных вычислительных устройств, которые могут собирать и отправлять данные без участия человека.

Все устройства, которые могут подключаться к интернету имеющие сетевой адрес являются частью интернет вещей. Например, камера которая осуществляет видеонаблюдение.

Система интернет вещей состоит из датчиков устройств, средств подключения, инструментов обработки данных и пользовательского интерфейса. Взаимодействие устройств в системе интернет вещей осуществляется через облачное соединение. Датчики устройства собирают данные в определенной среде и после сбора данных устройство отправляет их в облако. Это может быть сделано по *Wi-Fi*, *Bluetooth* или путем прямого подключения к интернету через *Ethernet*. Как только данные попадают в облако, они обрабатываются программным обеспечением для принятия решения о выполнении определенных действий. Эти шаги могут включать отправку предупреждений или автоматическую настройку датчиков устройства без или с участием пользователя.

Широкое распространение получила система «Умный дом» включающая в себя различную бытовую кухонную технику, музыкальные системы, интеллектуальные системы освещения, электрические жалюзи, автоматические окна и двери. Система используется для автоматизации определенных задач и обычно управляется удаленно. Сегодня число бытовых устройств, подключенных к сети исчисляется миллионами, и оно стремительно растет.

Больше данные в автомобильной промышленности. В настоящее время перспективнейшим направлением использования больших данных является автомобильная промышленность. Уже сейчас в автомобилестроении большие данные играют огромную роль, позволяя решать ряд вопросов: от маркетинга и продаж до прогнозирования неисправностей и обеспечения безопасности водителя. Все больше корпораций пытаются подключить автомобиль в сеть. Благодаря тому, что многие современные автомобили используют передовые технологии, предусматривающие многочисленные сенсоры, встроенные вычислительные инструменты и процессоры подключение их к сети не составляет больших проблем, а применение *Wi-Fi* позволило в машинах использовать прямое подключение к интернету для навигации по *GPS*, доступа к электронной почте и прослушивания музыки онлайн. Через сеть можно обновлять и изменять имеющиеся настройки блоков управления автомобилей. В подключенных автомобилях интернет послужит гарантией своевременного обновления всех приложений и передачей информации о состоянии авто в базу данных. Поскольку большие данные – это информация, собранная из многочисленных источников данных, на основании их анализа можно сделать вывод и о состоянии водителя.

По мере роста на рынке автомобилей с возможностями беспилотного управления, большие данные будут увеличиваться и накапливаться. И это в свою очередь служит основой для дальнейшего прогресса в сфере автомобильной промышленности так как чем больше накопленных данных, тем больше потенциал для автономных технологий.

Один из примеров сервиса по работе с большими автомобильными данными – платформа *Remoto*. Это проект дистанционного управления функциями автомобиля, его адаптацию с популярными авто брендами и масштабирование технологической платформы, которая позволит поддерживать несколько миллионов одновременно подключенных транспортных средств. На сегодняшний день в мире насчитывается уже более 500 тысяч мобильных пользователей *Remoto* [3].

Remoto состоит из управляющего модуля для автомобиля, серверного программного обеспечения в облаке *Microsoft* и мобильного приложения для дистанционного управления функциями машины. Среди ключевых возможностей *Remoto* можно отметить возможность удаленно управлять климатом внутри автомобиля, центральным замком, определять местоположение автомобиля и маршрут его передвижения, получать уведомление в случае его повреждения или эвакуации, получать информацию об эффективном использовании топлива, анализировать стиль вождения и получать подсказки по его улучшению. Несомненно, данная платформа крайне интересна как для коммерческого транспорта, так и для частных лиц. Доступный уже сегодня инструментарий открыл пути для принципиально нового взаимодействия с автовладельцем [3].

Заключение. Интернет вещей – это впечатляющий прорыв в мире технологий, который вносит существенные изменения в различные сферы нашей жизни. Это не только новая эпоха развития технологий, но и возможность создать более умный и устойчивый мир.

Список литературы

- [1] https://trends.rbc.ru/trends/innovation/5d6c020b9a7947a740fea65c#card_5d6c020b9a7947a740fea65c_1 (дата обращения: 14.02.2024).
[2] <https://www.kaspersky.ru/resource-center/definitions/what-is-iot> (дата обращения: 14.02.2024).
[3] <https://habr.com/ru/companies/bright-box/articles/325718/> (дата обращения: 14.02.2024).

Авторский вклад

Авторы внести равнозначный вклад

BIG DATA IN THE AUTOMOTIVE INDUSTRY

A.A.Sasnovsky

*Master, senior lecturer at the
Department of Communications,
Military Faculty of BSUIR*

R.A.Bozhko

*Master, cycle manager
Department of Military Communications
Faculty of BSUIR*

Abstract. Annotation. This article examines the significance of using big data in various fields.

Keywords: big data, information technology, internet of things.

УДК 621.38:64.06

УСТРОЙСТВО УПРАВЛЕНИЯ ИСПОЛНИТЕЛЬНЫМИ МЕХАНИЗМАМИ НА МИКРОКОНТРОЛЛЕРЕ ДЛЯ УМНОГО ДОМА



Я.В. Лисовецкий
Студент кафедры теоретических основ электротехники БГУИР,
lisovetskiy@bk.ru



Т.Н. Дворникова
Старший преподаватель кафедры информационных радиотехнологий БГУИР,
магистр технических наук
dvornikova@bsuir.by

Я.В. Лисовецкий

Студент кафедры теоретических основ электротехники БГУИР. Область научных интересов связана с программированием микроконтроллеров.

Т.Н. Дворникова

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с программированием встраиваемых систем, исследованием быстрого преобразования корректирующих кодов.

Аннотация. На базе микроконтроллера *ATMEGA* разработано устройство управления исполнительными механизмами для регулирования, поддержания и измерения температуры.

Микроконтроллеры широко применяются в различных областях науки и техники, в частности для управления различными исполнительными механизмами. В процессе решения поставленных задач использованы принципы системного подхода, теория схмотехнического проектирования радиоэлектронных средств, аналитические и физико-математические методы, методы компьютерной обработки данных.

Разработанное устройство управления исполнительными механизмами на микроконтроллере для умного дома виде макета может быть использовано в промышленных помещениях, школах, учебных заведениях, как решение в системах умного дома.

Введение. Внедрение новых технологий в современных системах, позволит автоматизировать различные виды устройств, используемых как в промышленности, так и в быту. В частности, это позволит пользователям управлять и контролировать работу устройств дистанционно и автоматически.

В настоящее время наиболее актуальным и перспективным направлением развития технологий современных систем является технология *IoT (Internet of Things)*, которая нашла свое применение в комплексных системах управления, таких как «умный дом».

Возлагать на человека задачу контроля и регулирования температуры технологических процессов в эпоху высоких компьютерных технологий просто не рационально. Сегодня все эти задачи возложены на микропроцессорные устройства.

Таким образом, задача разработки устройств с использованием микроконтроллеров для системы умного дома является актуальной задачей.

Описание принципа работы устройства. Устройство управления состоит из трех функциональных блоков и четырех внешних периферийных устройств, соединенных между собой.

Микроконтроллер *ATMEGA328P*, который является основным блоком платы расширения, имеет расширенный набор линий ввода/вывода общего назначения и периферийных устройств, подключённых к двум улучшенным периферийным шинам, две высокопроизводительные шины и 32-битную матрицу, состоящую из множества шин. Микроконтроллер *ATMEGA328P* имеет 8-канальный 10-разрядный АЦП в корпусах *TQFP* и *QFN/MLF*, два 8-битных Таймера\Счетчика с Отдельным Пред делителем частоты и Режимом сравнения, счетчик реального времени с отдельным генератором шесть ШИМ-каналов, также он имеет большое количество внешних интерфейсов, таких как *SPI*, *UART*, *I2C*. На рисунке 1 представлено модуль отладочной платы *ATMEGA328P*.

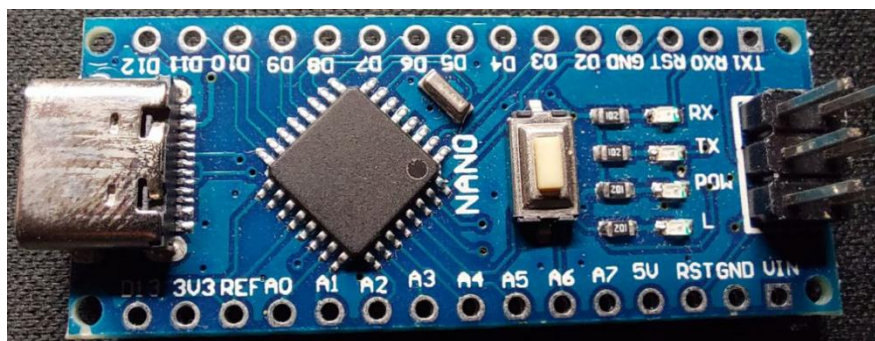


Рисунок 1. Модуль отладочной платы *ATMEGA328P*

Модуль *I2C* для управления *LCD* дисплеем (2×16 / 20×4) с помощью *ATMEGA328P*. Данный модуль позволяет уменьшить количество используемых выводов контроллера, вместо 8 или 4-битного соединения, требуется только 2 вывода (*SDA* и *SCL*).

На рисунке 2 представлен модуль интерфейса *I2C*.



Рисунок 2. Модуль интерфейса *I2C*

Интерфейс *I2C* – последовательная асимметричная шина. Он используется для подключения датчиков и периферийных устройств.

Модуль, построен он на микросхеме *PCF8574T*. Резисторы *R8* (4.7кОм) и *R9* (4.7кОм) необходимы для подтяжки линий *SDA* и *SCL*, в идеале при подключении двух и более устройств по шине *I2C* необходимо использовать подтяжку только на одном устройстве, позже напишу почему. На плате предусмотрены три переключки (по схеме видно, что линии *A0*, *A1*, *A2* подтянуты к питанию через резисторы *R4*, *R5*, *R6*), необходимы они для смены адресации устройства, всего их 8 вариантов. Изменение адресации дает нам возможность подключения до восьми устройств по шине *I2C* с микросхемой *PCF8574T*, варианты адресов показаны на рисунке (по умолчанию адрес устройства 0x27). Так же модуль оснащен потенциометром *R11* с его помощью можно изменить контрастность *LCD* дисплея. На рисунке 3 представлен модуль *LCD 1602+I2C BLUE*.



Рисунок 3. Модуль *LCD 1602+I2C BLUE*

Интерфейсы для программирования устройства управления исполнительными механизмами на базе микроконтроллера *ATMEGA328P*. На рисунке 4 представлено расположение выводов отладочной платы относительно разъемов, а также интерфейсы взаимодействия.

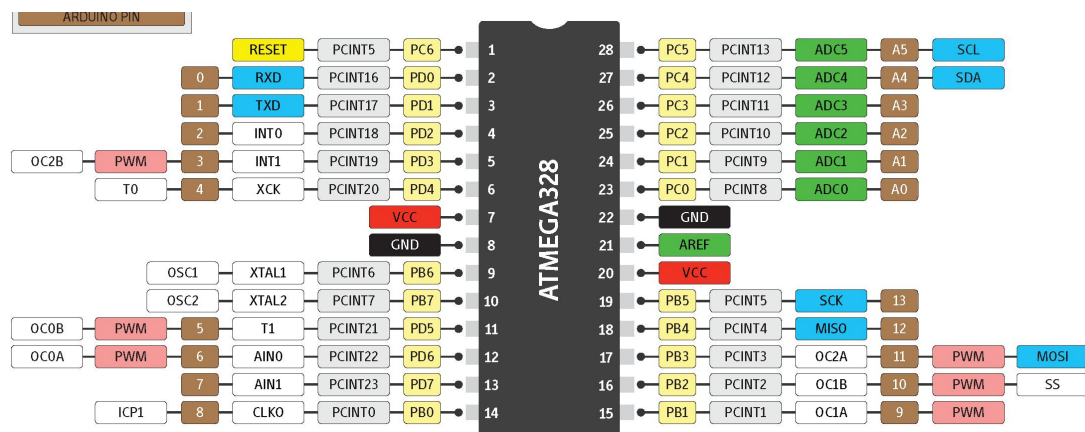


Рисунок 4. Расположение выводов и интерфейсы в отладочной плате на базе *STM32*

Метод программирования последовательных конфигурационных постоянных запоминающих устройств, который осуществляется через интерфейс *JTAG*, что упрощает

процесс отладки. Программа микроконтроллера, разработанная в программном обеспечении на персональном компьютере, поступает через интерфейс *mini-USB*, после чего программатор подключается к микроконтроллеру на плате, либо через *SWD*-интерфейс к иному устройству. Для программирования и отладки по интерфейсу *SWD* используются следующие выводы микроконтроллера:

- *SWDIO* – *Serial Wire Data Input/Output*;
- *SWCLK* – *Serial Wire Clock*;

Линии интерфейса имеют внутреннюю подтяжку, поэтому нет необходимости устанавливать внешние резисторы.

Модуль *LCD 1602+I2C BLUE* подключены напрямую к информационным выводам микроконтроллера.

Кнопка сброса представляет собой кнопку очистки памяти. При необходимости ее можно перепрограммировать на иное использование.

Для взаимодействия светодиодной индикации и кнопки сброса с контроллером используется интерфейс ввода/вывода общего назначения (*general-purpose input/output, GPIO*), который осуществляет прямое подключение компонентов к контроллеру.

Интерфейс *I2C* – последовательная асимметричная шина. Он используется для подключения датчиков и периферийных устройств.

Программа, разработанная в программном обеспечении на персональном компьютере, поступает через интерфейс *mini-USB*, после чего преобразуется в переходнике *USB-UART*, который осуществляет преобразование компьютерного интерфейса *USB* в последовательный интерфейс *UART*.

UART (Universal Asynchronous Receiver-Transmitter) – асинхронный последовательный интерфейс. У контроллера *ESP8266* два модуля *UART*. Один из которых устанавливает связь с отладочной платой на базе микроконтроллера *ATMEGA328P* по шине *UART*, а другой – для загрузки кода по последовательной связи.

SPI (Serial Peripheral Interface) – последовательный периферийный интерфейс. *ESP8266* имеет два *SPI (SPI и HSPI)* в ведущем и подчиненном режимах.

Разработка программной части устройства управления. Для разработки устройства управления исполнительными механизмами на микроконтроллере для умного дома используются аппаратные и программные средства. В процессе создания этапы проектирования и разработки неотделимы друг от друга. Решение конечной задачи достигается путем нахождения компромисса в выборе проектной платформы и среды разработки.

Проектные платформы, использующиеся для разработки устройства управления исполнительными механизмами на микроконтроллере для умного дома, определяют процесс проектирования и разработки. Широко используемыми в производстве платформами можно назвать промышленные персональные компьютеры, микроконтроллеры, сигнальные процессоры, программно-аппаратные комплексы и контроллеры с программируемой логикой, программируемые логические интегральные схемы и сверхбольшие интегральные схемы программируемой логики.

Одним из важнейших критериев при разработке устройства управления исполнительными механизмами на микроконтроллере для умного дома является выбор среды программирования.

Программирование микроконтроллеров *AVR* на языке *C* является одной из наиболее популярных и распространенных методик разработки встраиваемых систем. *AVR (Advanced Virtual RISC)* – это семейство 8-битных микроконтроллеров, разработанных компанией *Atmel* и используемых в различных применениях, включая индустрию, автомобильную отрасль, медицинскую технику и другие.

Основным преимуществом использования микроконтроллеров AVR является их простота в программировании и широкий набор интегрированных периферийных устройств. Большинство микроконтроллеров AVR имеют встроенную память программ и данных, а также различные модули, такие как АЦП, UART, SPI, I2C и другие, что делает их идеальным выбором для разработки различных электронных устройств.

Один из основных языков программирования, используемых для разработки на микроконтроллерах AVR, – это язык C. C – это язык высокого уровня, который обеспечивает простоту и гибкость в написании программ для AVR. Он позволяет разработчикам использовать все доступные функции и возможности микроконтроллера AVR, такие как GPIO, таймеры, прерывания и т. д.

Основные принципы программирования микроконтроллеров AVR на языке C включают в себя объявление и инициализацию переменных, использование операторов условия и циклов, работу с памятью, обработку прерываний и другие. Код на языке C затем компилируется с использованием специального компилятора, который превращает его в машинный код, понятный микроконтроллеру AVR.

Существует множество инструментов и сред разработки для программирования микроконтроллеров AVR на языке C, таких как Atmel Studio, AVR-GCC, AVRDUDE и другие. Эти инструменты предоставляют разработчикам удобную среду с отладчиками, эмуляторами, симуляторами и другими функциями, которые помогают в отладке и тестировании программного обеспечения для AVR.

Несмотря на свою простоту, программирование микроконтроллеров AVR на языке C требует от разработчиков хорошего понимания аппаратной архитектуры микроконтроллеров, умения работать с регистрами, битовыми манипуляциями.

Заключение. Данное устройство является актуальным, так как, с помощью разработанного устройства управления исполнительными механизмами на микроконтроллере для умного дома, решаются конкретные задачи пользователей. Данное устройство может применяться не только для умного дома, но и внедряться в производство.

Разработанное устройство обладает преимуществом относительно существующих решений, таким как:

- относительный режим работы;
- возможность управления устройством через Bluetooth и WiFi.

Имеется возможность модернизации и улучшения функций. Данное устройство интегрируется с другими системами умного дома.

Список литературы

- [1] Керниган Брайан У. Язык программирования C / Ритчи Деннис М., Керниган Брайан У. – М. : Вильямс, 2017. – 288 с.
- [2] Марц Н, Уоррен Д. Большие данные. Принципы и практика построения масштабируемых систем обработки данных в реальном времени. М.: Вильямс, 2016. – 368 с.
- [3] Тейлор Дж, Рэйден Т. Получение конкурентных преимуществ путем автоматизации принятия скрытых решений. М.: Символ-плюс, 2009. – 448 с.

Авторский вклад

Дворникова Татьяна Николаевна – постановка задачи разработки, руководство разработкой устройства управления исполнительными механизмами на микроконтроллере для умного дома.

Лисовецкий Ярослав Владимирович – написание программного кода, формирование структуры статьи.

MICROCONTROLLER-BASED ACTUATOR CONTROL DEVICE FOR SMART HOMES

U.V. Lisovetsky

*Student at the Department of
Theoretical Foundations of
Electrical Engineering, BSUIR*

T.N. Dvornikova

*Senior lecturer of the Information
Radio Technologies Department of
BSUIR, Master of Technical
Sciences*

Annotation. Based on the ATMEGA microcontroller, an actuator control device has been developed for regulating, maintaining and measuring temperature. Microcontrollers are widely used in various fields of science and technology, in particular to control various actuators. In the process of solving the assigned problems, the principles of the systems approach, the theory of circuit design of radio-electronic equipment, analytical and physical and mathematical methods, and methods of computer data processing were used. The developed device for controlling actuators on a microcontroller for a smart home in the form of a prototype can be used in industrial premises, schools, educational institutions, as a solution in smart home systems.

УДК 378

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ СМЕШАННОГО ОБУЧЕНИЯ В УЧЕБНОМ ПРОЦЕССЕ УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ КАК СПОСОБ ПОВЫШЕНИЯ УДОВЛЕТВОРЕННОСТИ СТУДЕНТОВ



Т.М. Фильченкова

старший преподаватель кафедры защиты информации БГУИР, аспирант кафедры психологии и педагогического мастерства Республиканского института высшей школы
t.filchenkova@bsuir.by

Т.М. Фильченкова

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой технологии управления качеством смешанного обучения в учреждениях высшего образования в условиях цифровой трансформации.

Аннотация. Рассмотрена организация смешанного обучения на примере преподавания учебной дисциплины студентам ИТ-специальностей факультета информационной безопасности учреждения образования «Белорусский государственный университет информатики и радиоэлектроники». Представлены обобщённые данные, полученные в результате опроса 128 студентов, которые изучали в весеннем семестре 2022/2023 учебного года дисциплину «Основы машинного обучения». На базе собранных сведений выявлено какие компетенции по мнению самих студентов развились благодаря применению технологии смешанного обучения, представлены результаты их анализа, сформулированы выводы.

Ключевые слова: информационно-коммуникационные технологии, образовательный процесс, смешанное обучение, цифровая трансформация образования.

Введение. Американские исследователи Дж. Бейли, Дж. Бергман, К. Бонк, Ч. Грэхем, Дж. Руни, Х. Стейкер, А. Сэмз, М. Хорн в своих опубликованных работах [1–5] представили виды смешанного обучения, существующие эффективные практики и методические подходы по организации данного обучения.

Изучению технологии смешанного обучения в преподавании учебных дисциплин и цифровой трансформации образования посвящен ряд работ белорусских и российских исследователей, в том числе В.А. Богуша [6], Ю.И. Капустина [7], А.А. Ковалевской, В.Н. Курбацкого, Н.В. Ломоносовой [8], Л.Г. Титаренко и др.

Применение технологии смешанного обучения в высшем образовании способствует его инновационному развитию. В настоящее время повысить эффективность обучения можно благодаря внедрению в образовательный процесс информационно-коммуникационных технологий.

Подготовку кадров по ИТ-специальностям удобно осуществлять по технологии смешанного обучения, т.к. процесс обучения в таком случае представляет собой модель будущей профессиональной деятельности, которая связана с удалённым взаимодействием.

Смешанное обучение – это образовательная технология, сочетающая традиционное и электронное обучение. Каждая из составляющих смешанного обучения важна в образовательном процессе. Электронное обучение связано с получением знаний с помощью информационно-коммуникационных технологий и различных электронных устройств (компьютеров, планшетов и смартфонов).

В Республике Беларусь в учреждениях высшего образования электронное обучение организовано в модульно объектно-ориентированной динамической обучающей среде Moodle, которая относится к классу систем управления обучением *LMS (Learning Management System*, с англ. – система управления обучением) [9]. Данное программное обеспечение позволяет организовывать дистанционное обучение.

Таким образом, *Moodle LMS* представляет собой систему управления ресурсами электронного обучения.

Исследователи считают, что цифровая трансформация образования затрагивает содержание, методы, организацию учебного процесса в динамично развивающейся обучающей среде [6]. Образовательный процесс по технологии смешанного обучения имеет меньшую трудоёмкость, т.к. часть очных занятий переведена в онлайн-формат и на самостоятельную управляемую работу студентов.

С 2020 года в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» (БГУИР) начался экспериментальный проект «Апробация смешанной модели обучения по ИТ-специальностям в рамках трансформации БГУИР в «Цифровой университет». Этот проект рассчитан на 5 лет.

Цель исследования, которое изложено в данной статье, – определить проблемы применения технологии смешанного обучения в преподавании дисциплин для студентов ИТ-специальностей на основе опроса студентов БГУИР.

Данное исследование соответствует приоритетным направлениям, обозначенным в Концепции развития системы образования в Республике Беларусь до 2030 года и Концепции цифровой трансформации процессов в системе образования Республики Беларусь на 2019-2025 годы.

В 2022/2023 учебном году преподаватели кафедры высшей математики использовали технологию смешанного обучения при преподавании учебной дисциплины «Основы машинного обучения». Данный курс является самостоятельной математической дисциплиной несмотря на то, что находится на стыке прикладной статистики, численных методов оптимизации, дискретного анализа [10].

В 2022/2023 учебном году в весеннем семестре преподаватели кафедры высшей математики использовали технологию смешанного обучения при преподавании учебной дисциплины «Основы машинного обучения» для студентов 2-го курса факультета информационной безопасности специальностей: 1-45 01 01-01 «Инфокоммуникационные технологии (системы телекоммуникаций)», 1-45 01 01-02 «Инфокоммуникационные технологии (сети инфокоммуникаций)», 1-45 01 01-04 «Инфокоммуникационные технологии (цифровое теле- и радиовещание)», 1-45 01 01-05 «Инфокоммуникационные технологии (системы распределения мультимедийной информации)», 1-98 01 02 «Защита информации в телекоммуникациях». 50% лекционных и практических занятий проводились дистанционно в асинхронном режиме в системе *Moodle LMS*.

Учебная дисциплина «Основы машинного обучения» состоит из 5 разделов. По данному курсу разработан электронный образовательный ресурс, который состоит из 5 модулей по разделам. В каждом модуле содержится текстовый материал с теоретическими сведениями и подробно разобранными примерами, а также тесты для проверки уровня усвоения студентами определенных математических понятий и их свойств.

Тесты приведены после каждой изучаемой темы, а их 11 тем. Тестирование позволяет студентам, изучающим курс «Основы машинного обучения» осуществлять самоконтроль, повторение и осмысление учебного материала, а преподавателям – корректировать свою работу со студентами в течение семестра.

Анкетирование студентов ИТ-специальностей БГУИР проводился на базе авторского опросника, представленного в *Google*-форме октябре – ноябре 2023 года. Поскольку объектом выполняемого диссертационного исследования выступает качество высшего образования в условиях цифровой трансформации, а предметом выполняемого диссертационного исследования – управление качеством смешанного обучения в учреждениях высшего образования, в анкете присутствовали вопросы, связанные с оценкой самими респондентами результатов их усвоения материала учебной дисциплины, которая преподавалась по технологии смешанного обучения.

Представим полученные данные, результаты их анализа и выводы. Характеристика выборки. В опросе приняли участие 128 студентов 3 курса факультета информационной безопасности следующих ИТ-специальностей:

- 19 студентов специальности 1-45 01 01-01 «Инфокоммуникационные технологии (системы телекоммуникаций)»,
- 23 студента специальности 1-45 01 01-02 «Инфокоммуникационные технологии (сети инфокоммуникаций)»,
- 20 студентов специальности 1-45 01 01-04 «Инфокоммуникационные технологии (цифровое теле- и радиовещание)»,
- 26 студентов специальности 1-45 01 01-05 «Инфокоммуникационные технологии (системы распределения мультимедийной информации)»,
- 40 студентов специальности 1-98 01 02 «Защита информации в телекоммуникациях».

Всего студентов этих специальностей обучалось на факультете информационной безопасности в весеннем семестре 2022/2023 учебного года 180 человек.

Результаты анкетирования студентов. В анкете респондентам был задан вопрос со шкалой от 0 (совсем не понравился) до 10 (очень понравился) баллов «Вам понравился смешанный формат изучения дисциплины «Основы машинного обучения?». Результаты оценивания применения технологии смешанного обучения представлены на рисунке 1.



Рисунок 1. Гистограмма ответов студентов факультета информационной безопасности УО «БГУИР»

На вопрос о том насколько понравился смешанный формат изучения дисциплины опрошенные студенты прокомментировали свою оценку (по желанию):

1 10 и 9 баллов оценили, потому что удобно проходить курс, доступные материалы в электронном образовательном ресурсе, чередование форматов занятий делало их не такими скучными, не надо тратить время на дорогу, комфортно изучать материал в своем темпе, в таком формате удобнее было планировать свое время, дома можно было успеть сделать задание быстрее чем за пару, за меньшее потраченное время на выходе получился такой же результат знаний;

2 8 и 7 баллов оценили, т.к. не все идеально, мало практических занятий, довольно интересно, но хотелось бы больше примеров в обучающих материалах для практических и лабораторных задач (на Python), есть нюансы в практической части (не всегда удаётся в полной мере реализовать такой вид обучения);

3 6 и 5 баллов – потому что было мало практики, только теория, слишком быстрое прохождение материала, необычный формат обучения;

4 4, 3 и 2 балла – потому что было скучно, не все понятно.

Студентам был задан вопрос «Что можно было улучшить, на ваш взгляд, в преподавании дисциплины «Основы машинного обучения» в смешанном формате?». Респонденты ответили:

- добавить дистанционные лекции,
- постараться подать материал более понятно,
- преподавательские компетенции.

Как видно из ответов на данный вопрос, студенты хотели бы, чтобы преподаватели усовершенствовали свои компетенции. В [11, с. 447] отмечено, что в современных условиях высшего образования преподавателям необходимо учитывать разнообразие обучающихся и их индивидуальные потребности, быть гибкими и адаптироваться к различным стилям обучения, уровням знаний и способностям студентов, чтобы обеспечить качественный и инклюзивный образовательный процесс. Это всё в свою очередь требует от преподавателей большего вложения энергии и ресурсов. Респондентам был задан вопрос «Как вы считаете преподаватели больше сил и труда затратили на преподавание дисциплины «Основы машинного обучения по технологии смешанного обучения?». На рисунке 2 представлена усреднённая инфографика ответов на данный вопрос.



Рисунок 2. Круговая диаграмма ответов студентов факультета информационной безопасности УО «БГУИР» на вопрос об энергоресурсности преподавателей при работе в смешанном формате

Анализируя ответы на вопрос об энергоресурсности преподавателей при работе в смешанном формате, можно сделать вывод: большинство студенты, принявших участие в опросе считают, что преподаватели затратили больше сил и труда при работе по технологии смешанного обучения.

Важно было узнать у респондентов ответ на вопрос «Вы до изучения дисциплины «Основы машинного обучения» имели опыт обучения по педагогической технологии смешанного обучения?» для более точного анализа результатов исследования. На рисунке 3 представлена круговая диаграмма ответов студентов факультета информационной безопасности УО «БГУИР» на вопрос об опыте обучения по технологии смешанного обучения.

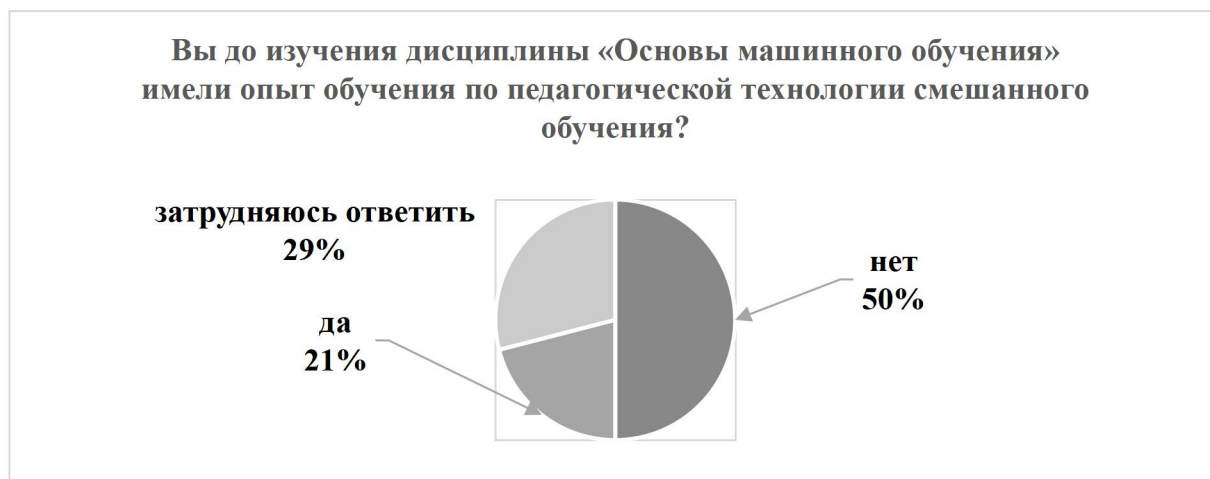


Рисунок 3. Круговая диаграмма ответов студентов факультета информационной безопасности УО «БГУИР» на вопрос об опыте обучения по технологии смешанного обучения

Только 21 % студентов, участвующих в опросе, имели опыт обучения по технологии смешанного обучения. Таким образом, большинство студентов ранее не обучались в таком формате. Следовательно, у них было больше трудностей с изучением дисциплины «Основы машинного обучения». Для выявления этого обстоятельства студентам был задан вопрос «Вам пришлось больше заниматься самостоятельно при смешанном обучении, чем при традиционном аудиторном формате?». На рисунке 4 показана круговая диаграмма ответов студентов факультета информационной безопасности УО «БГУИР» на вопрос о самостоятельной работе при смешанном обучении.

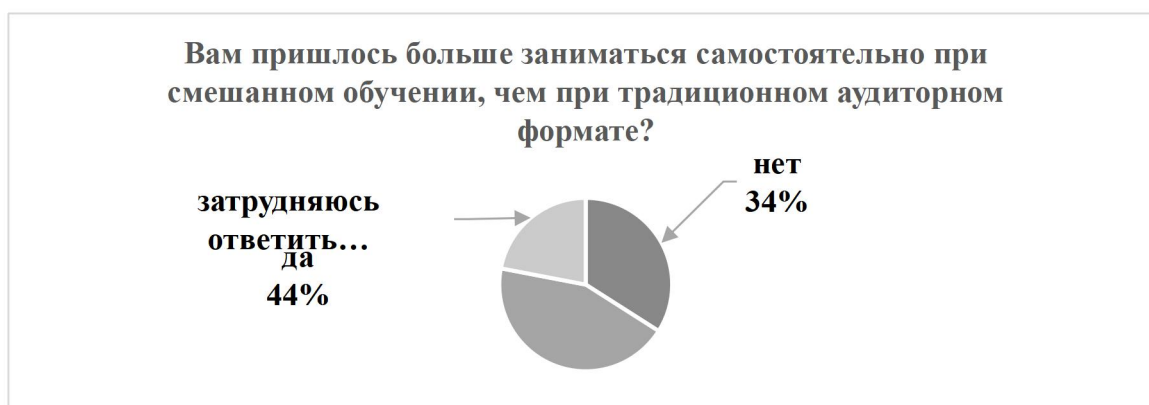


Рисунок 4. Круговая диаграмма ответов студентов факультета информационной безопасности УО «БГУИР» на вопрос о самостоятельной работе при смешанном обучении

Как выше было сказано, качество высшего образования зависит от интеллектуального потенциала студентов, показателем которого можно считать успеваемость студентов. Респондентам был задан вопрос «Вы довольны результатом сдачи экзамена по дисциплине «Основы машинного обучения»?».

На рисунке 5 представлена круговая диаграмма ответов студентов факультета информационной безопасности УО «БГУИР» на вопрос о результатах сдачи экзамена по дисциплине, которая преподавалась в смешанном формате.



Рисунок 5. Круговая диаграмма ответов студентов факультета информационной безопасности УО «БГУИР» на вопрос о результатах сдачи экзамена по дисциплине, которая преподавалась в смешанном формате

Большинство студентов, принявших участие в данном опросе, довольны результатами сдачи экзамена по дисциплине, которая преподавалась по технологии смешанного обучения. Количественная оценка успеваемости студентов факультета информационной безопасности УО «БГУИР» представлена на рисунке 6.

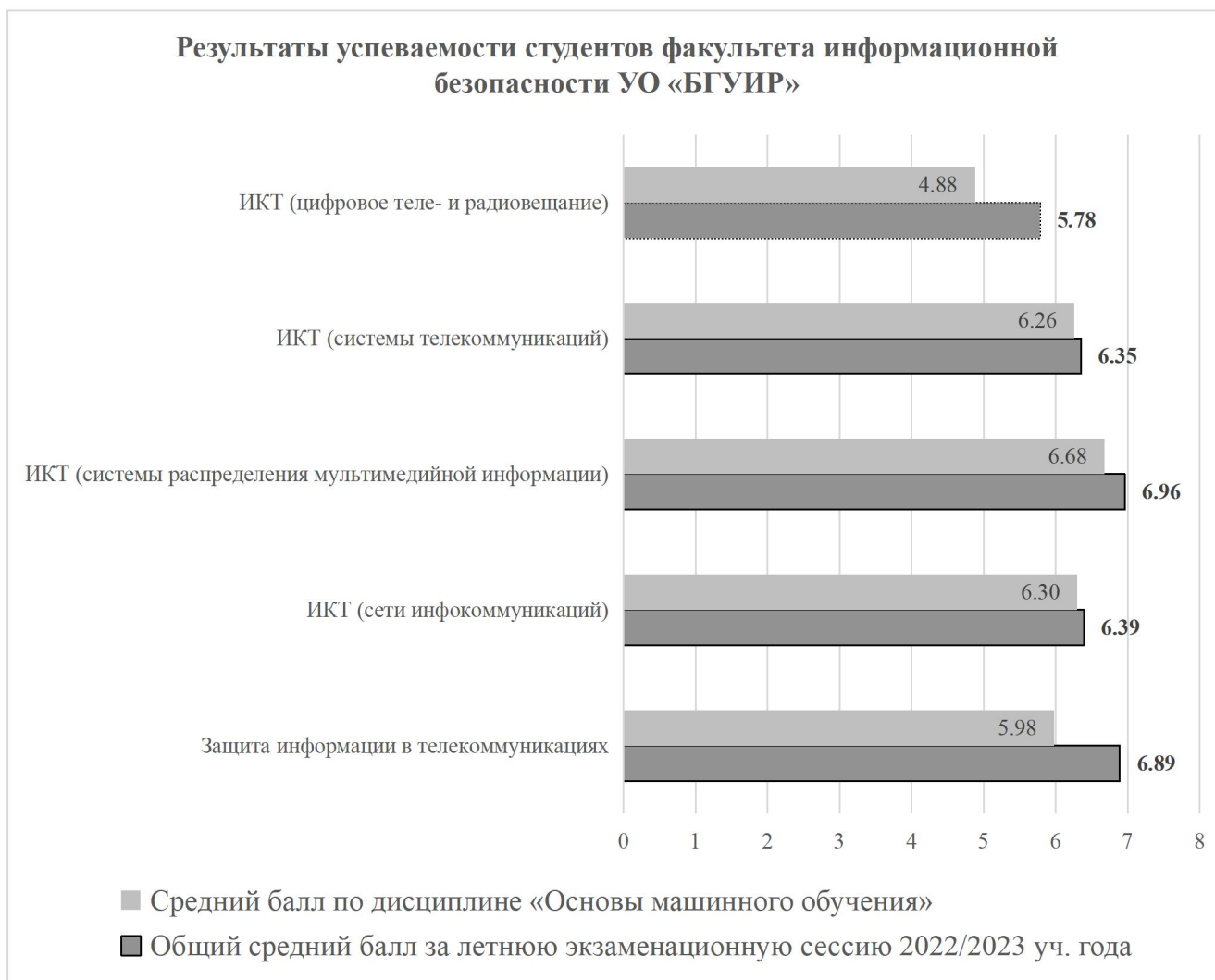


Рисунок 6. Линейчатая гистограмма результатов сдачи летней экзаменационной сессии 2022/2023 учебного года студентами 2 курса 5 специальностей факультета информационной безопасности УО «БГУИР»

Как видно из рисунка 6, лучше всего сдали экзамен по дисциплине «Основы машинного обучения» студенты специальности Инфокоммуникационные технологии (системы распределения мультимедийной информации), средний балл составил 6,68. У студентов этой специальности и общий средний балл по результатам сдачи летней экзаменационной сессии 2022/2023 учебного года выше.

Рассмотрим подробнее результаты сдачи экзамена по дисциплине «Основы машинного обучения» студентами ИТ-специальностей факультета информационной безопасности УО «БГУИР» (рис. 7 и 8).

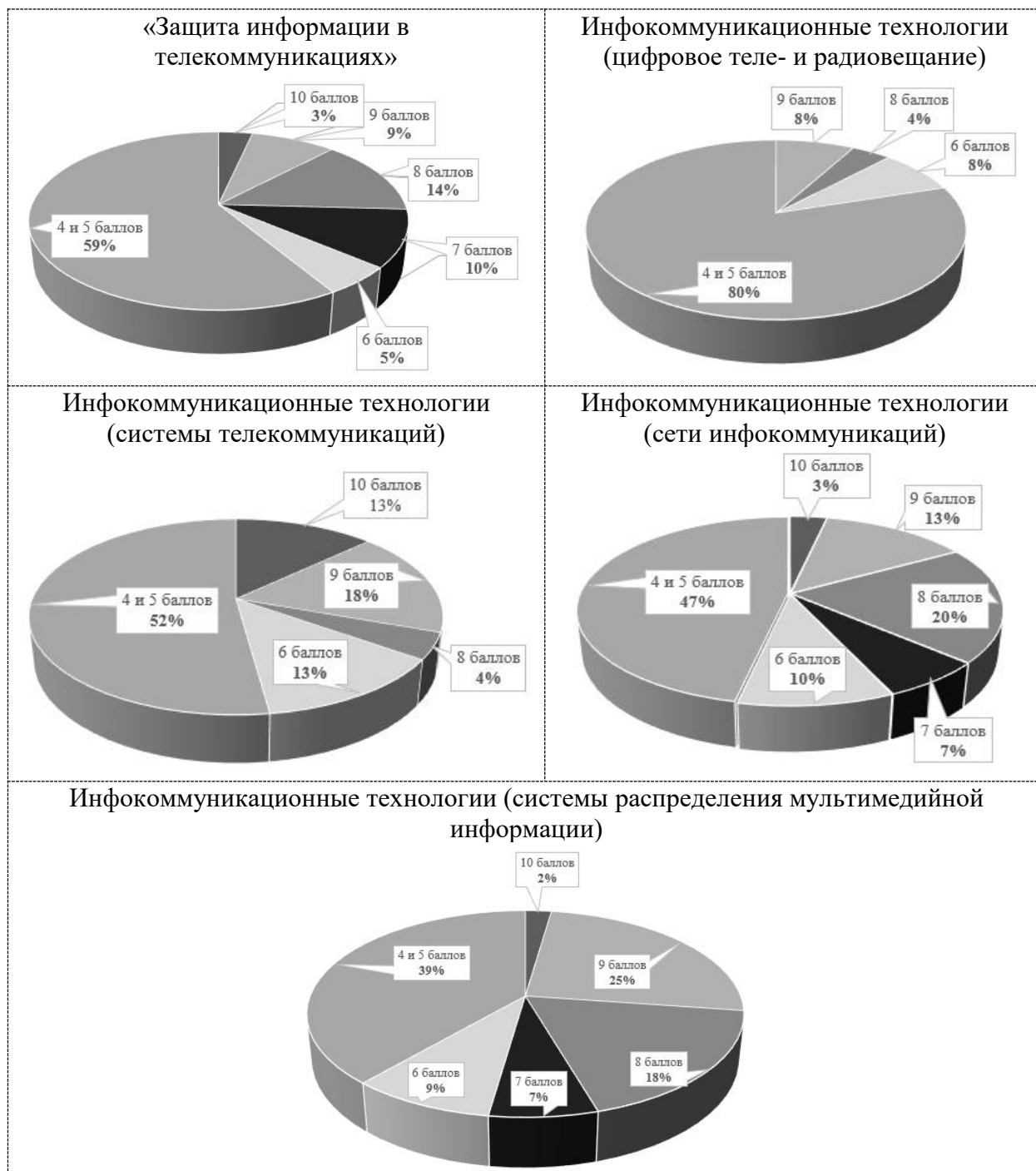


Рисунок 7. Результаты сдачи экзамена по дисциплине «Основы машинного обучения» студентами ИТ-специальностей факультета информационной безопасности УО «БГУИР»

обучения» по технологии смешанного обучения: самостоятельная работа, самообучение, тайм-менеджмент, самодисциплина.

Как отмечено в [12], эффективное решение первостепенной задачи, которая решается системой образования в условиях удалённого взаимодействия, будет в том случае, если оно учитывает актуальный опыт самостоятельной работы студентов, который приобретён ими в ходе образовательного процесса, организованного удалённо. Роль самостоятельной работы при смешанном обучении важна.

Проведён корреляционный анализ ответов студентов факультета информационной безопасности УО «БГУИР». В результате выявлены следующие прямые взаимосвязи умеренной значимости:

– между высокой оценкой технологии смешанного обучения и хорошей успеваемостью (выше 5,0 баллов);

– между наличием опыта обучения по технологии смешанного обучения и высокой успеваемостью (выше 8,0 баллов).

Выводы. Проведённое исследование по выявлению проблем в образовательном процессе, организованном с применением технологии смешанного обучения в преподавании дисциплин для студентов ИТ-специальностей, позволило выявить следующее:

– смешанное обучение предполагает большую самостоятельную работу студента;

– электронный образовательный ресурс учебной дисциплины должен быть хорошо разработан;

– преподаватели должны обладать сформированной компетенцией, направленной на использование дистанционных образовательных технологий в своей профессиональной деятельности.

Технология смешанного обучения является перспективной, но требует от преподавателей и студентов инновационного развития:

– профессорско-преподавательский состав учреждений высшего образования должен на высоком уровне разрабатывать дидактические ресурсы для эффективной самостоятельной работы студентов в электронной образовательной среде;

– за удалённой работой студентов в электронной образовательной среде, а именно контролировать успеваемости, необходимо следить деканатам и кураторам учебных групп для выявления проблемных моментов в учебном процессе по технологии смешанного обучения и их дальнейшему решению, что позволит улучшить качество высшего образования.

Список литературы

[1] Rooney, J. E. Blending learning opportunities to enhance educational programming and meetings / J. E. Rooney // Association Management. – 2003. – Vol. 55. №. 5. – P. 26–32.

[2] Bonk, C.J. The Handbook of Blended Learning: Global Perspectives, local designs / C. J. Bonk, C. R. Graham – San Francisco, CA : Pfeifer. – 2006.

[3] Staker, H., Horn M.B. Classifying K-12 Blended Learning / H. Staker, M. B. Horn [Электронный ресурс]. – Режим доступа: <http://www.christenseninstitute.org/wp-content/uploads/2013/04/Classifying-K-12-blended-learning.pdf>. – Дата доступа: 30.11.2023.

[4] Bergmann, J. Flip your classroom: Reach every student in every class every day / J. Bergmann, A. Sams. – International Society for Technology in Education. – 2012.

[5] Bailey, J. Navigating the Digital Shift: Implementation Strategies for Blended and Online Learning / J. Bailey, C. Schneider, T.V. Ark. – Digital Learning Now! Smart Series. – 2013.

[6] Богуш, В. А. Цифровизация образования: проблемы, вызовы и перспективы / В. А. Богуш, Е. Н. Шнейдеров // Адукацыя і выхаванне. – 2021. – № 1. – С. 14–21.

[7] Капустин, Ю. И. Педагогические и организационные условия эффективного сочетания очного обучения и применения технологий дистанционного образования: Автореф. дис. ... д-ра пед. наук. М. : РАО, 2007. – 40 с.

[8] Ломоносова, Н. В. Система смешанного обучения в условиях информатизации высшего образования: Дис. ... канд. пед. наук. М. : МПГУ, 2018. – 191 с.

[9] Шнейдеров, Е. Н. Система мониторинга показателей образовательного процесса для образовательных платформ на базе Moodle LMS / Е. Н. Шнейдеров [и др.] // BIG DATA и анализ высокого уровня = BIG DATA and Advanced Analytics : сборник научных статей IX Международной научно-практической конференции, Минск, 17–18 мая 2023 г. : в 2 ч. Ч. 2 / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. А. Богуш [и др.]. – Минск, 2023. – С. 188-194.

[10] Князюк, Н. В. Некоторые аспекты преподавания дисциплины «Основы машинного обучения» / Н. В. Князюк, О. В. Рыкова // Научные и методические аспекты математической подготовки в университетах технического профиля : материалы V Международной научно-практической конференции, Гомель, 27 апреля 2023 г. / Белорусский государственный университет транспорта ; ред. Ю. И. Кулаженко. – Гомель, 2023. – С. 139 – 141.

[11] Пунчик, В. Н. Образовательный брендинг преподавателя как технология диверсификации в условиях цифровой среды // Диверсификация педагогического образования в условиях развития информационного общества : материалы II Междунар. науч.-практ. конференции, Минск, 16 ноября 2023 г. / БГУ, Каф. педагогики и проблем развития образования ; [редкол.: Г. В. Пальчик (гл. ред.) и др.]. – Минск : БГУ, 2023. – С. 446–450.

[12] Пунчик, В. Н. Трансформация самостоятельной учебной деятельности учащихся в условиях удалённо организованного образовательного процесса / В.Н. Пунчик, М.С. Подолякина // Адукацыя і выхаванне. 2021. – № 2 (350). – С. 68–79.

USING BLENDED LEARNING TECHNOLOGY IN THE EDUCATIONAL PROCESS OF A HIGHER EDUCATION INSTITUTION AS A WAY TO INCREASE STUDENT SATISFACTION

T.M. Fichenkova

*Senior lecturer of the Department
of Information Security of BSUIR,
postgraduate student of
the Department of Psychology and
Pedagogical Excellence of
State Institution of Education
«National Institute for Higher
Education»*

Abstract. The article considers the organisation of blended learning on the example of teaching the academic discipline to the students of IT-specialties of the Faculty of Information Security of the educational institution «Belarusian State University of Informatics and Radioelectronics». The generalised data obtained as a result of the survey of 128 students who studied in the spring semester of the academic year 2022/2023 the discipline «Fundamentals of Machine Learning». On the basis of the collected data it is revealed what competences the students themselves think have been developed due to the application of blended learning technology, the results of their analysis are presented, conclusions are formulated.

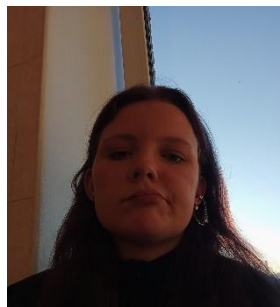
Keywords: information and communication technologies, educational process, blended learning, digital transformation of education.

УДК 004.032.26

НЕЙРОННЫЕ СЕТИ ДЛЯ ОБНАРУЖЕНИЯ ЛИЦ НА ИЗОБРАЖЕНИЯХ



Кабариха В.А.
Ассистент кафедры
инженерной психологии и
эргономики
v.kabarikha@bsuir.by



Семерник О.Д.
Студент третьего курса
факультета
компьютерного
проектирования БГУИР
semernik2135@gmail.com

В.А.Кабариха

Окончил Белорусский государственный университет. Область научных интересов связана с компьютерными системами, компьютерными сетями, интерактивными разработками, современными языками программирования.

О.Д. Семерник

Студент третьего курса Белорусского государственного университета информатики и радиоэлектроники. Область научных интересов связана с разработкой бизнес требований, создание информационных систем и разработкой программного обеспечения для решения бизнес-задач.

Аннотация. Определение лиц в изображениях – это важная задача в области компьютерного зрения. ВВ статье рассмотрены как классические методы, так и современные подходы к определению лиц с использованием нейронных сетей.

Показано, что могут быть разработаны алгоритмы распознавания лиц на изображениях в различных условиях с высокой степенью точности.

Ключевые слова: определение лиц, нейронные сети, каскадные классификаторы, компьютерное зрение.

Введение. Определение лиц в изображениях – это важная задача в области компьютерного зрения. Системы, способные точно обнаруживать и локализовывать лица, имеют широкий спектр применений, начиная от безопасности и видеонаблюдения до автоматической обработки фотографий и видео.

Определение лиц позволяет идентифицировать человека на изображении. Это может быть полезно для систем контроля доступа, аутентификации или поиска пропавших людей.

Распознавание лиц также может помочь определить эмоциональное состояние человека, что может быть полезно в маркетинге, психологии и медицине. Определение лиц позволяет автоматически тегировать фотографии в социальных сетях или создавать удобные фотоальбомы.

Использование нейронных сетей для обнаружения лиц на изображениях повышает эффективность и делает возможным создание систем распознавания лиц людей в реальном времени. Нейронные сети помогают точно локализовать лица и

идентифицировать людей, что заключается в определении границы лица на изображении и сравнении лица с базой данных для идентификации конкретных личностей.

Нейронные сети продолжают совершенствоваться, и будущее связано с повышением точности определения лиц, что может быть достигнуто путём развития методов, способных работать с разными условиями освещения, позы и шумом. С развитием вычислительных мощностей и с ростом возможностей аппаратных платформ нейронные сети будут все более эффективно применяться в реальном времени для обработки видео и фотографий.

Классические методы определения лиц. Определение лиц в изображениях – это важная задача в области компьютерного зрения. За последние десятилетия было предложено множество методов для обнаружения и распознавания лиц. Были рассмотрены следующие методы определения лиц и объектов на изображении:

1 Метод главных компонент (*PCA*) – это статистический метод, который используется для снижения размерности данных. В контексте определения лиц, *PCA* применяется для выделения наиболее информативных признаков изображения. Основная идея заключается в том, чтобы найти новые оси (главные компоненты), вдоль которых разброс данных максимален. Эти оси позволяют представить изображение в пространстве меньшей размерности.

2 Линейный дискриминантный анализ (*LDA*) – это метод, который также используется для снижения размерности данных, но с учетом классов. Он стремится максимизировать разделение между классами. В контексте определения лиц, *LDA* помогает выделить признаки, которые наилучшим образом разделяют лица разных людей.

3 Метод независимых компонент (*ICA*) – это метод, который позволяет разложить смешанный сигнал на независимые компоненты. В случае определения лиц, *ICA* может помочь выделить различные части лица (например, глаза, нос, рот) как независимые компоненты.

4 Скрытые марковские модели (*HMM*) – это статистическая модель, которая используется для моделирования последовательностей данных. В определении лиц, *HMM* может помочь учесть динамические изменения в изображении лица.

5 Метод упругой связи графов (*EBGM*) – это метод, который использует графовую структуру лица для определения лиц. Он учитывает связи между различными частями лица.

6 Метод опорных векторов (*SVM*) – это метод машинного обучения, который используется для классификации данных. В определении лиц, *SVM* может помочь разделить лица на разные классы (например, разные люди).

7 Интегральный дискриминантный анализ (*IDA*) – это метод, который комбинирует преимущества *PCA* и *LDA*. Он позволяет учесть, как общую информацию о данных, так и разделение между классами.

Эти методы были успешно применены до появления глубокого обучения и остаются важными в исследованиях и практических приложениях. Однако с развитием нейронных сетей, таких как сверточные нейронные сети (*CNN*), были достигнуты еще более высокие результаты в определении лиц.

Использование сверточных нейронных сетей в определении лиц на изображении. Сверточные нейронные сети (*CNN*) – это мощный инструмент в области компьютерного зрения, который демонстрирует выдающиеся результаты в задачах обнаружения и распознавания лиц. При рассмотрении сверточных нейронных сетей было выделено следующее:

- архитектура сверточных нейронных сетей;
- обучение на больших объемах данных;
- инвариантность к позе и освещению;

- применение в реальном времени;

Архитектура сверточных нейронных сетей (*CNN*) состоит из нескольких слоев, включая: сверточные слои, слои объединения (пулинга) и полносвязные слои. Сверточные слои обрабатывают изображение, выделяя важные признаки, такие как грани, текстуры и формы. *CNN* обучаются на больших наборах данных, содержащих изображения лиц разных людей и автоматически извлекают признаки, которые помогают различать лица.

CNN обладают свойством инвариантности к некоторым трансформациям, таким как повороты, масштабирование и изменение освещения. Это позволяет успешно определять лица в разных условиях. *CNN* могут работать в реальном времени, что важно для систем видеонаблюдения, автоматической фотоиндексации и других приложений.

Faster R-CNN – эта архитектура объединяет сверточные слои с регионами интереса (*ROI*) для обнаружения лиц. *MTCNN (Multi-task Cascade Convolutional Neural Network)* – этот метод использует несколько сверточных сетей для обнаружения и локализации лиц на разных масштабах.

Алгоритм и пример работы *Multi-task Cascade Convolutional Neural Network*. *MTCNN (Multi-task Cascade Convolutional Neural Network)* – это алгоритм, разработанный для обнаружения и локализации лиц на изображениях. При рассмотрении алгоритма были выделены следующие аспекты работы:

- архитектура *MTCNN* состоит из нескольких сверточных нейронных сетей, объединенных в каскад. каждая сеть выполняет определенную задачу: обнаружение лица, локализацию лица и точное определение ключевых точек (например, глаза, нос, рот).

- *MTCNN* показывает высокую точность обнаружения лиц на изображении и способность работать с разными ориентациями, уровнями освещения и другими факторами.

Для работы с *MTCNN* использовалась библиотека *MTCNN* на *Python*. Пример программы, выполняющей поиск лица на рисунке 1:

```
# Загрузка изображения
image = cv2.imread('C:/flutter/photo_131@29-03-2021_19-56-51.jpg')

# Преобразование изображения в оттенки серого и сохранение его на диск
gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

cv2.imwrite('C:/flutter/photo_131@29-03-2021_19-56-51-face-gray.jpg', gray)

# Обнаружение лица с помощью MTCNN
faces = detector.detect_faces(image)
draw_image_with_boxes('C:/flutter/photo_131@29-03-2021_19-56-51-face.jpg', faces)
```

Рисунок 1. Пример программы для определения лиц на изображениях

На рисунке 2 представлено исходное изображение для поиска лица. Лицо на изображении имеет небольшой поворот, что усложнит задачу точного определения границ лица.



Рисунок 2. Изображения для обнаружения лица

Для уменьшения размерности изображения за счёт снижения количества каналов и увеличения за счёт этого скорости обработки изображения, его необходимо привести в оттенки серого. Это также повышает устойчивость к освещению и снижает уровень шума на изображении. На рисунке 3 изображено исходное изображение в оттенках серого.



Рисунок 3. Изображение в оттенках серого для обнаружения лица

После того, как код программы отработал, результат с обнаруженными лицами записывается в отдельный файл. На рисунке 4 показано изображение с обнаруженными лицами.



Рисунок 4. Изображение обнаруженного лица

MTCNN используется в системах видеонаблюдения, автоматической фотоиндексации, аутентификации и других приложениях.

Заключение. Классические методы, такие как *PCA*, *LDA* и *ICA*, были успешно применены до появления глубокого обучения и сверточных нейронных сетей. Однако классические методы используют созданные вручную признаки, в то время как сверточные нейронные сети автоматически извлекают признаки из данных и могут обучаться на больших объемах данных, что позволяет им адаптироваться к различным условиям.

С развитием сверточных нейронных сетей (*CNN*) были достигнуты высокие результаты в определении лиц на изображениях и будущее связано с повышением точности и надежности определения лиц. Это особенно важно для систем безопасности и видеонаблюдения. Сверточные нейронные сети могут успешно применяться в реальном времени для обработки видео и фотографий. Сверточные нейронные сети смогут работать в различных условиях освещения, поз человека, с различным уровнем шума на изображении и в условиях других факторов, которые снижают точность определения лиц.

Список литературы

- [1] Ali, W., Tian, W., Din, S.U. et al. Classical and modern face recognition approaches: a complete review. *Multimed Tools Appl* 80, 4825–4880 (2021). <https://doi.org/10.1007/s11042-020-09850-1>
- [2] Mliki, H., Dammak, S. & Fendri, E. An improved multi-scale face detection using convolutional neural network. *SIViP* 14, 1345–1353 (2020). <https://doi.org/10.1007/s11760-020-01680-w>
- [3] How to Perform Face Detection with Deep Learning [Электронный ресурс]. – Режим доступа: <https://machinelearningmastery.com/how-to-perform-face-detection-with-classical-and-deep-learning-methods-in-python-with-keras/>. – Дата доступа: 12.02.2024.

Авторский вклад

Кабариха Виктор Александрович – руководство исследованием нейронных сетей для обнаружения лиц на изображении.

Семерник Ольга Дмитриевна – тестирование нейронных сетей для поиска лиц, исследование актуальности и перспектив в использовании нейронных сетей, формирование структуры статьи.

NEURAL NETWORKS FOR FACE RECOGNITION IN IMAGES

V.A. Kabariha
Assistant department
of Engineering
Psychology and
Ergonomics, BSUIR

O.D. Semernik
Third-year student
at the Faculty of
Computer Design, BSUIR

Abstract. Detecting faces in images is an important task in the field of computer vision. The article discusses both classical methods and modern approaches to face detection using neural networks.

It is shown that algorithms for recognizing faces in images can be developed with high accuracy under various conditions.

Keywords: Face detection, neural networks, cascade classifiers, computer vision.

UDC 004.021:004.75

WEIBULL DISTRIBUTION BASED MODEL BEHAVIOR IN COLOR INVARIANT SPACE FOR BLIND IMAGE QUALITY EVALUATION



A.M. Gavrovska

Assistant Professor, Department of Telecommunications at the University of Belgrade - School of Electrical Engineering, Serbia, PhD in Electrical Engineering and Computer Science, anaga777@gmail.com, anaga777@etf.bg.ac.rs



A.B. Samčović

Full Professor of Information and Communication Technologies at the University of Belgrade - Faculty of Transport and Traffic Engineering, Serbia, PhD andrej@sf.bg.ac.rs



D.M. Dujković

IT&HR Manager at the University of Belgrade - School of Electrical Engineering, Serbia, M.Sc. in Electrical Engineering dragi@etf.bg.ac.rs



Y.I. Golub

Associate Professor, United Institute of Informatics Problems National Academy of Sciences of Belarus (UIIP NAS), PhD in Engineering Sciences, 6423506@gmail.com



V.V. Starovoitov

Chief Researcher, United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Doctor of Engineering Sciences, Professor, valerys@newman.bas-net.by

A.M. Gavrovska

Graduated from the University of Belgrade - School of Electrical Engineering, Serbia (Dipl.Ing), and acquired PhD in electrical engineering and computer science at the University of Belgrade - School of Electrical Engineering. Current position is Assistant Professor. The area of scientific interests is related to the development of methods and algorithms in the field of signal and image processing, multimedia, video and telemedicine technologies, information and communication systems.

A.B. Samčović

Andreja B. Samčović is currently a Professor of Information and Communication Technologies at the University of Belgrade – Faculty of Transport and Traffic Engineering. He received his Dipl.Ing and PhD degrees at the University of Belgrade. He is a visiting professor and scholar worldwide (Germany, Austria, England, Slovakia, Scotland, Romania...), where he taught lectures, seminars or tutorials. His research interest includes

signal and image processing, coding and compression, information security and information and communication technologies.

D.M. Dujković

Graduated from the University of Belgrade - School of Electrical Engineering, Serbia, and received his BSc., and MSc. at the Department for synthesis and analysis of electric circuits. The area of scientific interests is related to RF electronics, digital signal and image processing, linear and nonlinear analyzing methods, and algorithm development. He is currently IT&HR manager at the University of Belgrade - School of Electrical Engineering, Serbia.

Y.I. Golub

Senior Researcher United Institute of Informatics Problems National Academy of Sciences of Belarus (UIIP NAS). Graduated from the Belarusian National Technical University (2005). Post-Graduate Courses – UIIP NAS, 2006-2009, thesis titled «Algorithms for human recognition by iris images». The area of scientific interests is related to the development of methods and algorithms for processing and analyzing digital images.

V.V. Starovoitov

Graduated from the Belarusian State University. Laureate of the State Prize of the Republic of Belarus (2002). The area of scientific interests is related to the digital image analysis.

Abstract. In this paper, we analyze Weibull distribution based fitting using phase congruency components obtained via color invariance approach. Since such model is found suitable for image quality evaluations, we test the effects on different distorted images. Blind or no-reference image quality assessment enables testing without having available subjective ratings. Thus, it is possible to consider additional type of distortions usually neglected in available datasets through data augmentation artificial intelligence tools. The obtained results show that Weibull based modelling using color invariance domain components may be suitable for further experimental analysis. Moreover, data augmentation not used in the training phase, demonstrate the importance of development modern quality evaluator that can be applied even for modified data obtained as a result of augmentation.

Keywords: Probability distribution, fitting, distortion types, image quality assessment, data augmentation.

Introduction. When comparing the effectiveness of different quality evaluations, three approaches are considered, as well as appropriate fine tuning of the parameters. Namely, in addition to subjective image quality assessment (IQA) and personal ratings, there are: full-reference (FR), reduced-reference (RR) and non-reference (NR) evaluation approaches. FR metrics for IQA are expected to give the most satisfying results compared to pristine or original images named as reference [1, 2]. RR approach exploits the attributes of the reference images [3], while the most valuable are NR ones [4, 5]. The purpose is to design a new method for automatic quality assessment of digital images of various types using selected features and low complexity model which do not rely on pristine image content. This is also called blind IQA since no prior information about reference is available.

Recently, there are many blind IQA models and frameworks developed [4, 6-9]. One of the most successful approaches is Naturalness Image Quality Evaluator (NIQE) based on natural scene statistics (NSS) [4, 8]. High correlation with subjective ratings is obtained by block or patch-wise estimations. Typical features are subtracted and contrast normalized (MSCN) coefficients, where local activity is measured and only active patches are treated using Gaussian based fitting. Some of the IQA metrics include supervised model based approaches [9]. Recently, some of the novel unsupervised machine learning models are introduced [6-7]. In [6] unsupervised blind IQA based on joint structure and NSS features, called JSN-NIQE, is employed using thirty eight features per block and per scale (blocks of 96×96 pixels and two scales meaning additionally blocks of 48×48 pixels size). Among MSCN and other statistical features, phase congruency mapping is employed within NIQE-like framework. Such mapping is applied for structural measurement in color invariance domain. Inspired by the results presented in [6], such calculations are tested using different distortion types.

Materials and methods. One of the most applicable benchmarking datasets for IQA is LIVE database [10] applied initially only for FR method development. Five different distortions

can be found such the noise, two compressions, blurring and fading. Here, we compare JSN-NIQE measurement results and Weibull fitting results for available distortions and additional images not included in the dataset, but obtained through data augmented approach. Even though, such data may or may not be considered as natural ones structural similarity exists, as well as of importance to observe whether such data may be considered through color invariance domain [11].

The color invariance is calculated using common RGB (Red-Green-Blue) color space, leading to three components: $C_1=0.06*R+0.63*G+0.27*B$, $C_2=0.3*R+0.04*G-0.35*B$ and $C_3=0.34*R-0.6*G+0.17*B$. These components represent the result of transforming color values to the Gaussian color model in RGB terms [11]. The color invariance is expected to be robust to material properties, as well as illumination change or any other similar imaging conditions. Weibull fitting is employed for each component in order to observe the behavior of the coefficients representing the local activity of image/patch. In [6] such color space is employed in order to form the feature vector. Structure and attention methods enable to exploit gradient or edge information. Discontinuity found within distorted images is expected to have a crucial impact on subjective based evaluation. By color invariant domain, it is possible to evaluate different types of edge strength while enabling robust content measurements regardless of object reflectance, illumination, noise, highlights, shadows, and similar.

Figure 1 shows an example of pristine image and the three components. They can be interpreted as edge strength measures in order to evaluate the image (or patch) content activity.

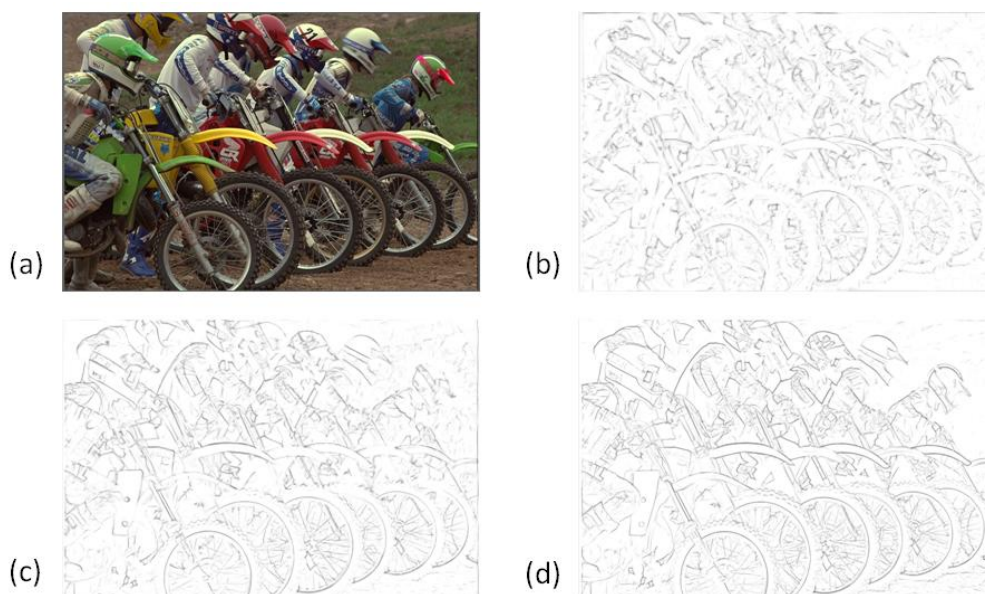


Figure 1. Example of (a) pristine image, and (b)-(d) three color invariance domain components.

Nowadays, the generative AI (artificial intelligence) is applied for content creation. Namely, models and libraries are developed for image generation. Data augmentation can be described as the process of producing new data from available content, primarily for overcoming data lack in learning phase. Even though, the augmented image content may produce biases in the learning step. It can be applied for further experimenting with test images as well, in order to analyze how possible random changes may affect the model, like in the case of quality assessment. There are many ways to perform data augmentation. Some of the typical examples are: by changing brightness, contrast, saturation, color spaces, by adding noise, or cropping, scaling, flipping, translation, rotation, etc. Currently, there are different tools that can be used for enabling data diversity [12]. Some of the Python libraries have been especially designed for this

task, like: Albumantations [13], Augmentor [14], Imgaug3 [15], Keras ImageDataGenerator [16] or some other similar tool. In the absence of available subjective scores intended for supervised machine learning evaluations, it is valuable to observe what the effects are obtained using novel generated data. Here, this is done for JSN-NIQE [6] and Weibull color invariance components. For defining data augmentation Albumantations library [13] is applied. In this paper, pipeline consists of horizontal flip and a random selection of brightness and contrast selection using [13].

Results. In order to quantify the differences obtained with augmented data compared to pristine and originally distorted image [12], the selected pipeline is applied primarily to generate additional image data. Depending on the distorted image type, local activity throughout content is changing, which also affects the patch-wise results used in blind measurements. In Figure 2 Jpeg compression distorted image and Fast fading based distorted image are processed with the pipeline for obtaining Data augmentation 1 and Data augmentation 2 samples, respectively. Moreover, obvious discontinuities are presented in the color invariant domain.

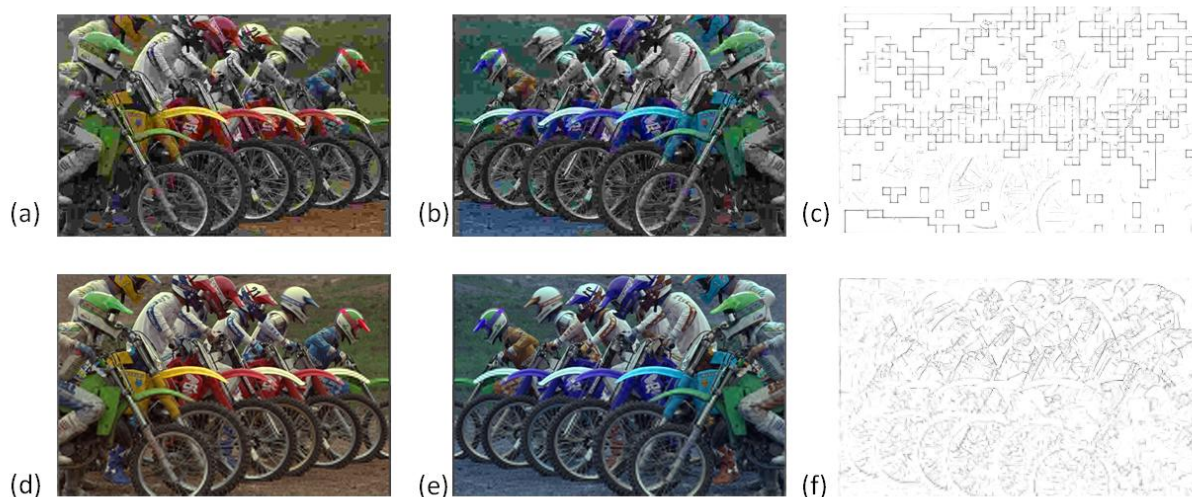


Figure 2. (a) Jpeg distorted image and corresponding (b) result of data augmentation and (c) C1 component; (d) Fast fading (ff) based distorted image and corresponding (e) result of data augmentation and (f) C2 component.

Phase congruency is used to map relevant discontinuities as dimensionless measure. For each color invariant component, histogram of obtained coefficients is approximated using two-parameter Weibull distribution. The two parameters are noted as a and b , representing scale and shape, respectively. In the case of jpeg compression distortion, for an image produced by augmentation fitting is performed globally here, and the comparison between components is presented in Figure 3.

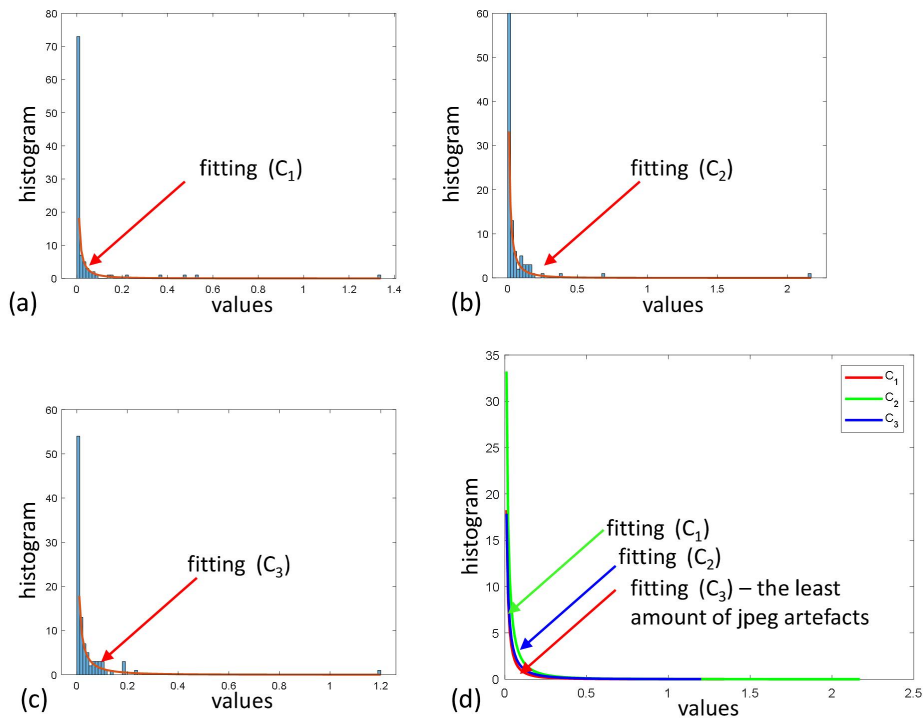


Figure 3. Fitting for (a) C_1 , (b) C_2 , and (c) C_3 component, and (d) comparison between the curves.

Table 1 shows the relative differences for the two parameters (Δa , Δb) in Data augmentation 1 for jpeg distorted images used as a source. The relative differences are calculated in terms of pristine image, showing C_2 component as the most robust here. On the other hand, absolute shape difference (b_{diff}) found between source and augmented version is less than difference calculated for the scale parameter, a_{diff} .

Table 1. Percentage of Weibull distribution parameter variations in Data augmentation 1 (jpeg)

Component	Δa [%]	Δb [%]	a_{diff} [%]	b_{diff} [%]
C_1	-24.5	-12.2	25.8	3.9
C_2	-2.5	0.1	2.5	-0.8
C_3	13.6	5.1	-27.1	-5.2

Table 2 shows the percentage of Weibull distribution parameter variations found in Data augmentation 2, where fast fading (ff) distorted image is treated as a source. Higher variations for fading are found expressed as the relative differences in terms of pristine reference. Nevertheless, the ff parameters are, in general, less affected compared to the jpeg distorted source.

Table 2. Percentage of Weibull distribution parameter variations in Data augmentation 2 (ff)

Component	Δa [%]	Δb [%]	a_{diff} [%]	b_{diff} [%]
C_1	32.5	26.4	-0.6	0.4
C_2	62.9	33.3	2.5	0.9
C_3	110	35.1	0.0	-0.2

Having in mind other features, like the NSS based ones, and JSN-NIQE measure, a comparison between measurements is given in Table 3. Distortion effects are evident in both distorted and augmentation cases. It is shown that unsupervised quality evaluation based on

augmented data may have different impact based on distortion type used as a source. Here, a decrease in score is obtained in the jpeg case, and metric increase for the fading.

Table 3. Unsupervised quality evaluation based on augmented data and distortion type.

<i>Data</i>	<i>Score</i>	<i>Relative difference [%]</i>
Original	4.3	/
Distorted 1 (jpeg)	15.8	267
Distorted 2 (ff)	12.9	200
Data augmentation 1 (jpeg)	15.5	260
Data augmentation 2 (ff)	13.2	207

In Figure 4 Weibull parameters are compared for both source distorted images and augmentation data. The parameters are found while distribution fitting for images, and are calculated globally, not for patches as in JSN-NIQE. It is illustrated that some of the parameters may perform weaker grouping, leading to more sparse clusters after the augmentation.

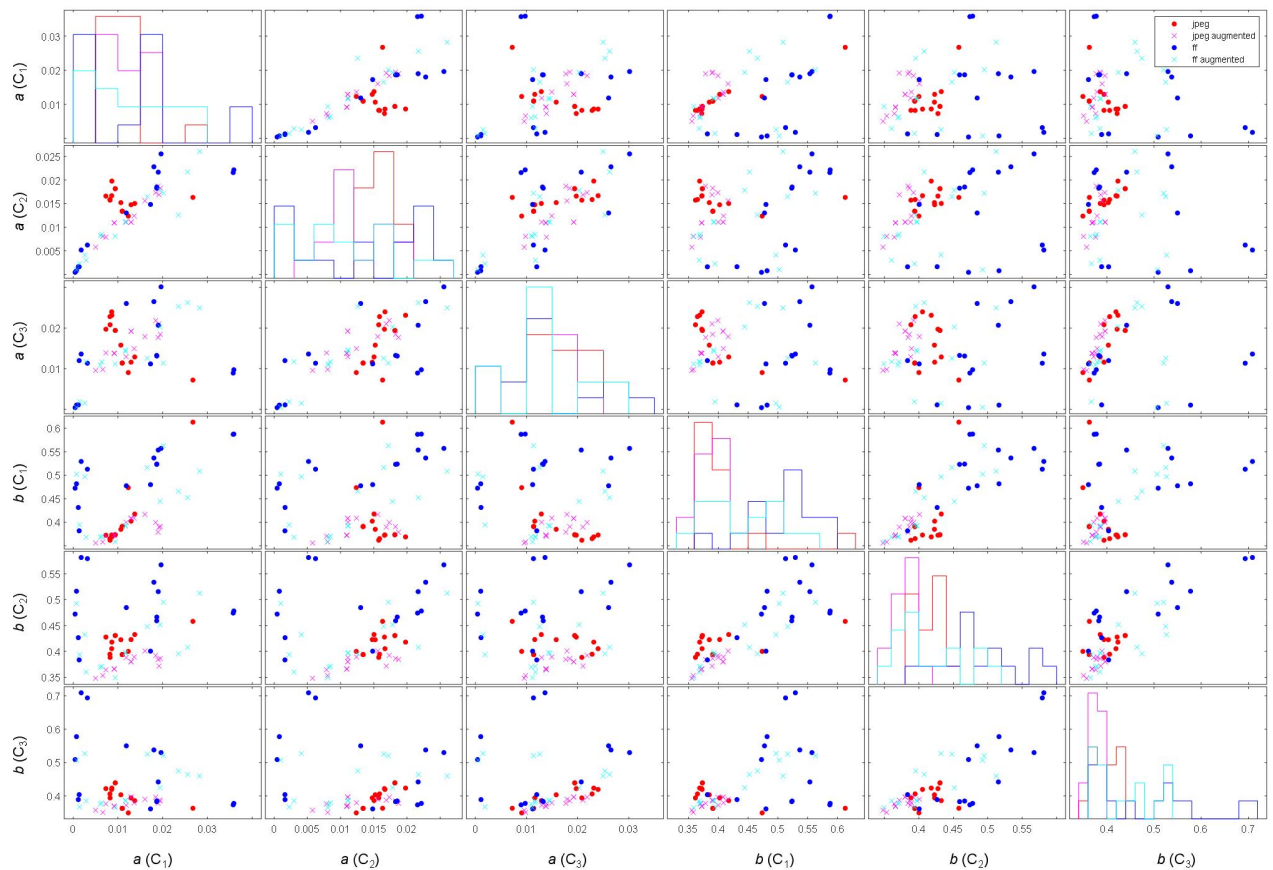


Figure 4. Weibull parameters comparison for jpeg and fast fading (ff) in the case of source and augmented data.

Conclusions. In this paper we analyze an unsupervised metric, and modification of Weibull parameters when applying a simple data augmentation procedure. Depending on the distortion source image, color invariant components are affected in different ways.

Weibull distribution is shown to be adequate for modelling task for each of the component. Even though, it should be noted that the fitting may not be suitable in some cases, like when the amount of distortion is extremely high. The estimation of Weibull parameters in each of the color invariant components represent an efficient and blind insight into existing distortions, and

seem to be valuable for further experimental analysis. It is expected that augmented data for big data tasks may be estimated in the similar manner.

Acknowledgements. The presented paper is a result obtained within bilateral cooperation with the Republic of Belarus supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia. This study is partially supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, 451-03-47/2023-01/200103.

Reference list

- [1] Rajchel, M. and Oszust, M. No-reference image quality assessment of authentically distorted images with global and local statistics. *Signal, Image and Video Processing*, 15, pp.83-91. 2021.
- [2] Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), pp.600-612, 2004.
- [3] Dost, S., Saud, F., Shabbir, M., Khan, M.G., Shahid, M. and Lovstrom, B. Reduced reference image and video quality assessments: review of methods. *EURASIP Journal on Image and Video Processing*, 2022(1), pp.1-31. 2022.
- [4] Mittal A., Soundararajan R., and Bovik A. C. Making a “completely blind” image quality analyzer. *IEEE Signal processing letters*, 20(3), pp. 209-212, 2012.
- [5] Zhao, K., Yuan, K., Sun, M., Li, M. and Wen, X. Quality-aware pre-trained models for blind image quality assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 22302-22313, 2023.
- [6] He, Q., Yang, C., Yang, F. and An, P. Unsupervised blind image quality assessment based on joint structure and natural scene statistics features. *Journal of Visual Communication and Image Representation*, 87, 2022. p.103579. <https://doi.org/10.1016/j.jvcir.2022.103579>
- [7] Saha, A., Mishra, S. and Bovik, A.C. Re-IQA: Unsupervised Learning for Image Quality Assessment in the Wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5846-5855. 2023.
- [8] Gavrovska, A., Dujković, D., Samčović, A., Golub, Y. and Starovoitov, V. No-reference local image quality evaluation. In *2022 30th Telecommunications Forum (TELFOR)*, pp. 1-4. IEEE. November, 2022.
- [9] Dendi, S.V.R. and Channappayya, S.S. No-reference video quality assessment using natural spatiotemporal scene statistics. *IEEE Transactions on Image Processing*, 29, pp.5612-5624. 2020.
- [10] Sheikh H.R., Sabir M.F. and Bovik A.C., A statistical evaluation of recent full reference image quality assessment algorithms, *IEEE Trans. on Image Processing*, 15(11), pp. 3440-3451, 2006
- [11] Geusebroek, J.M., Van den Boomgaard, R., Smeulders, A.W.M. and Geerts, H., Color invariance. *IEEE Transactions on Pattern analysis and machine intelligence*, 23(12), pp.1338-1350. Doi: 10.1109/34.977559. 2001.
- [12] Data augmentation, <https://research.aimultiple.com/data-augmentation-techniques/>
- [13] Albumentations, <https://pypi.org/project/albumentations/> (last accessed 10.02.2024.)
- [14] Augmentor, <https://pypi.org/search/?q=Augmentor> (last accessed 10.02.2024.)
- [15] Imgaug, <https://pypi.org/project/imgaug3/> (last accessed 10.02.2024.)
- [16] ImageDataGenerator, https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator (last accessed 10.02.2024.)
- [17] Kovési, P. Image features from phase congruency. *Videre: Journal of computer vision research*, 1(3), pp.1-26. 1999.

authors' contribution

Ana Gavrovska – led the research on evaluation the quality of tested images.

Andreja Samčović – statement of the research problem, description of the procedure.

Dragi Dujković – describing the data augmentation possibilities.

Yuliya Golub – analysis of quality assessment possibilities.

Valery Starovoitov – statement of the research problem, introducing distribution relevance.

ПОВЕДЕНИЕ МОДЕЛИ НА ОСНОВЕ РАСПРЕДЕЛЕНИЯ ВЕЙБУЛЛА В ЦВЕТОВО-ИНВАРИАНТНОМ ПРОСТРАНСТВЕ ДЛЯ СЛЕПОЙ ОЦЕНКИ КАЧЕСТВА ИЗОБРАЖЕНИЙ

А.М. Гавровска

*Доцент кафедры телекоммуникации
Белградского университета -
Электротехнически Факультет,
Сербия, доктор наук,*

А.В. Самчовић

*Профессор информационных и
коммуникационных
технологий Белградского
университета - Факультет
транспорта и транспортной
инженерии, Сербия, доктор
наук,*

Д.М. Дујковић

*Менеджер Белградского
университета -
Электротехнически
Факультет, Сербия,
кандидат наук,*

Ю.И. Голуб

*Старший научный сотрудник,
Объединенный институт проблем
информатики Национальной
академии наук Беларуси (ОИПИ
НАН), кандидат технических наук,*

В.В. Старовойтов

*Главный научный сотрудник
Объединенного института
проблем информатики
Национальной академии наук
Беларуси, доктор технических
наук, профессор,*

Аннотация. В этой статье мы анализируем аппроксимацию на основе распределения Вейбулла с использованием компонентов фазовой конгруэнтности, полученных с помощью подхода цветовой инвариантности. Поскольку такая модель признана подходящей для оценки качества изображения, мы проверяем ее влияние на различные искаженные изображения. Слепая оценка качества изображения или без эталонного изображения позволяет проводить тестирование без наличия субъективных оценок. Таким образом, можно учитывать дополнительный тип искажений, обычно игнорируемых в доступных наборах данных, с помощью инструментов искусственного интеллекта для увеличения данных. Полученные результаты показывают, что моделирование на основе Вейбулла с использованием компонентов области цветовой инвариантности может быть пригодным для дальнейшего экспериментального анализа. Более того, пополнение данных, не используемое на этапе обучения, демонстрирует важность разработки современного оценщика качества, который можно применять даже для модифицированных данных, полученных в результате пополнения.

Ключевые слова: Распределение вероятностей, аппроксимация, типы искажений, оценка качества изображения, увеличение данных.

UDC 004.021:004.75

BLOCKED ALGORITHM OF SHORTEST PATHS SEARCH IN SPARSE GRAPHS PARTITIONED INTO UNEQUALLY SIZED CLUSTERS



A.A. Prihozhy

*Professor at the Computer and System Software Department, Doctor of Technical Sciences, Full Professor
Belarusian National Technical University
prihozhy@yahoo.com*



O.N. Karasik

*Tech Lead at ISsoft Solutions (part of Coherent Solutions) in Minsk, Belarus, PhD in Technical Science
karasik.oleg.nikolaevich@gmail.com*

A.A. Prihozhy

Full professor at the Computer and system software department of Belarusian national technical university, doctor of science (1999) and full professor (2001). His research interests include programming and hardware description languages, parallelizing compilers, and computer aided design techniques and tools for software and hardware at logic, high and system levels, and for incompletely specified logical systems. He has over 300 publications in Eastern and Western Europe, USA and Canada. Such worldwide publishers as IEEE, Springer, Kluwer Academic Publishers, World Scientific and others have published his works.

O.N. Karasik

Tech Lead at ISsoft Solutions (part of Coherent Solutions) in Minsk, Belarus; PhD in Technical Science (2019). Interested in parallel computing on multi-core and multi-processor systems.

Abstract. *In this paper we consider the problem of searching shortest paths between all pairs of vertices of a directed weighted sparse graph which is partitioned into clusters by finding dense weakly connected subgraphs. We address this problem by developing new block-based algorithms that describe the shortest paths by matrices of blocks of unequal sizes corresponding to the sizes of the graph clusters. These algorithms extend the capabilities of known existing algorithms using blocks of equal size (such as the blocked algorithms of Floyd-Warshall family) with respect to adequate graph modeling of real networks of different purposes, and with respect to efficient use of parallelism and computational resources of multiprocessor systems and multi-core processors. The blocked algorithm of finding shortest paths in sparse large size graphs partitioned into clusters that is proposed in this paper reduces, on the one hand, the amount of memory used, and, on the other hand, reduces the number of block recalculations. Diagonal blocks describe shortest paths within clusters, non-diagonal compact blocks describe non numerous weighted arcs connecting clusters. Shortest paths between vertices of different clusters are computed in real time. The memory consumption is reduced compared to the Floyd-Warshall algorithm to a number of times equal to the number of clusters. In order to reduce the number of block recalculations, a new operation is introduced to accurately compute the shortest path between vertices of one cluster, passing through the vertices and edges of another cluster, as well as through the edges connecting the clusters. Applying this operation alone allows us to find solutions that introduce a small error (a few percent) in the lengths of the shortest paths when the weights of edges between clusters are small, and allows us to find exact solutions when the weights of these edges are increased. Accurate solutions can be obtained for sparse graphs modeling road, computer, and other networks.*

Keywords: *sparse graph, cluster, APSP problem, blocked algorithm, unequally sized blocks, heterogeneous system.*

Introduction. The problem of finding shortest and longest paths for all pairs of vertices in a large sparse weighted graph [1 – 6] has many application domains. Recently the emergence of

heterogeneous parallel computing systems [7 – 9] has increased interest in this problem. Many competitive algorithms are developed for various types of graphs and for different formulations of the problem: between two vertices; between the source (sink) and each other vertex (single source and single sink – SSSP); between each pair of vertices (all pairs shortest paths– APSP); to meet the requirement that all graph vertices must be listed on the path, etc.

The paper considers the APSP problem and algorithms [10, 11] of solving it. Two families of the APSP algorithms exist: 1) based on the Dijkstra SSSP-algorithm [1]; 2) based on the Floyd-Warshall APSP-algorithm [2]. The first family includes the Dijkstra algorithm [1], the Bellman-Ford algorithm [12], the Johnson algorithm [13], the Harish and Narayanan algorithm [14], and others [15]. The second family includes among others the Floyd-Warshall (*FW*) algorithm [2], the blocked Floyd-Warshall algorithm (*BFW*) proposed in [7, 10, 11] by Katz, Venkataraman and others, the graph extension-based algorithm (*GEA*) and the heterogeneous blocked APSP algorithm (*HBAPSP*) both proposed by Prihozhy and Karasik in [16 – 18]. The algorithms can be parallelised by OpenMP [19]. The results as follows have been obtained based on the idea of using blocks: a recursive blocked *FW* algorithm [10]; efficient usage of GPUs [7 – 9]; solving sparse graph scaling problem [20]; optimization of data allocation in hierarchical memory [21]; improving cache performance for APSP [11, 17, 22]; a cooperative threaded algorithm [23, 24]; selection of the optimal block-size [25]; reducing energy consumption [26]; search for shortest paths using dataflow networks of actors [27, 28]. In work [5], a method of inferring new blocked algorithms which divide the input graph into unequal subgraphs and divide the matrix of shortest path distances into blocks of unequal sizes has been proposed.

The key contribution of the paper is a fast and memory efficient blocked algorithm of computing the shortest paths within unequally sized clusters of a large sparse graph and computing the shortest paths between vertices of different clusters in real time.

Blocked all-pairs shortest paths algorithm for unequally sized blocks. Let $G = (V, E)$ be a simple directed graph with real edge-weights consisting of a set V , $|V| = N$ of vertices numbered 1 through N and a set E of edges. Let W be a cost adjacency matrix for G . So, $w(i, i) = 0$, $1 \leq i \leq N$; $w(i, j)$ is the cost (weight) of edge (i, j) if $(i, j) \in E$ and $w(i, j) = \infty$ if $i \neq j$ and $(i, j) \notin E$. Let d_{ij} be a length of shortest path from vertex i to vertex j , and D be a matrix of distances between all pairs of vertices $i, j \in V$, $i \neq j$. Let P be a matrix whose element p_{ij} is a vertex that is previous for vertex j in a path from i to j . The objective of an APSP-algorithm is to compute the D and P matrices for a given graph G .

In work [5], we have proposed to decompose the graph G into subgraphs (clusters) and decompose the matrix B into blocks of unequal sizes defined by vector $S = (S_1 \dots S_M)$ (Figure 1,a). While M blocks are square on the principal diagonal of B (block B_{ii} has the $V_i \times V_i$ size), all other blocks are rectangular in general case (block B_{ij} has the $V_i \times V_j$ size for $i, j = 1 \dots M$, $i \neq j$). All blocks in row i have the height of V_i , and all blocks in column j have the width of V_j . Matrix P of previous vertices in the shortest paths has the same structure.

At the aim of processing unequally sized clusters, we extended the known blocked Floyd-Warshall algorithm *BFW* to the all-pairs shortest path algorithm *BFWUS* [5], which can handle a block-matrix B of unequally sized blocks. *BFWUS* is described by Algorithm 1. In a loop along m that performs M iterations, it recalculates each of M^2 blocks of matrix B , therefore, it carries out M^3 recalculations in total. In terms of vertex count, the time complexity of *BFWUS* is N^3 and the memory complexity is N^2 since each block has the layout (Figure 1,b) of matrix of shortest path distances (DiM).

Figure 2 illustrates the operation of *BFWUS* and depicts the order of calculating blocks.

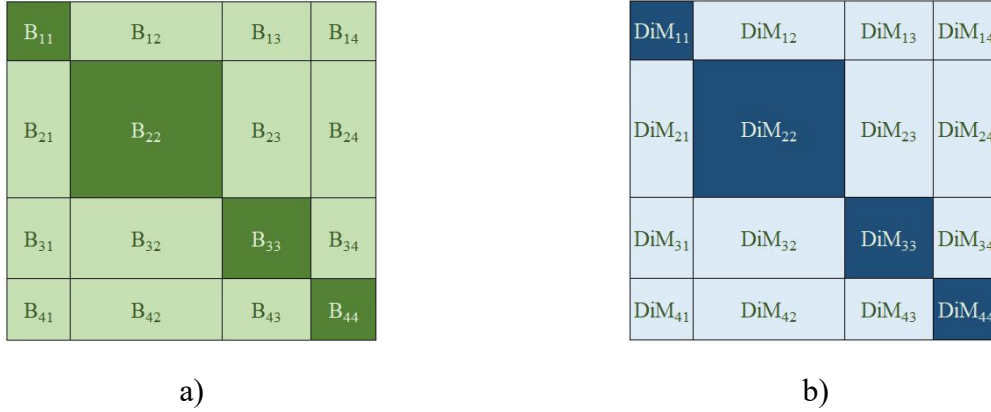


Figure 1. Blocked matrix B of shortest paths distances: a) diagonal blocks are square and non-diagonal blocks are rectangular; b) $BFWUS$ represents all blocks by matrix of shortest path distances (DiM)

Algorithm 1: Blocked APSP algorithm accounting for blocks of unequal sizes ($BFWUS$)

Input: A number N of vertices in input graph
Input: A matrix $W[N \times N]$ of graph edge weights
Input: A vector $S = (S_1 \dots S_M)$ of sizes of vertex subsets
Input: A number M of blocks per row (column)
Output: A blocked matrix $B[M \times M]$ of path distances
Output: A blocked matrix $P[M \times M]$ of previous vertices in shortest paths

```

for  $i, j \leftarrow 1$  to  $N$  do
    if  $W(i, j) \neq \infty$  then  $P^{init}(i, j) \leftarrow i$  else  $P^{init}(i, j) \leftarrow undefined$ 
 $B[M \times M] \leftarrow W[N \times N]$     $P[M \times M] \leftarrow P^{init}[N \times N]$ 
for  $m \leftarrow 1$  to  $M$  do
     $USBC(S, B, P, m, m, m)$  // D0
    for  $v \leftarrow 1$  to  $M$  do
        if  $v \neq m$  then
             $USBC(S, B, P, v, m, m)$  // C1
             $USBC(S, B, P, m, m, v)$  // C2
    for  $v \leftarrow 1$  to  $M$  do
        if  $v \neq m$  then
            for  $u \leftarrow 1$  to  $M$  do
                if  $u \neq m$  then
                     $USBC(S, B, P, v, m, u)$  // P3
return  $B, P$ 

```

Algorithm 2 describes a block-calculation procedure $USBC$ which has a feature of processing blocks of unequal sizes. The algorithm inputs are matrices B and P which describe blocks of sizes defined by vector S . Indices v, m and u choose in matrix B three blocks $B_{v,u}, B_{v,m}$ and $B_{m,u}$ of which two or three can be identical. The indices choose similar blocks in matrix P . The sizes of blocks are $S_v \times S_u, S_v \times S_m$ and $S_m \times S_u$ respectively. $USBC$ consists of three nested loops. It makes $S_m \cdot S_v \cdot S_u$ attempts to update the values of elements of blocks $B_{v,u}$ and $P_{v,u}$. The order of loops is essential. The loop along k must be the outer, it cannot be reordered with other loops.

The key difference *USBC* against Floyd-Warshall is that the non-diagonal blocks are rectangles but squares.

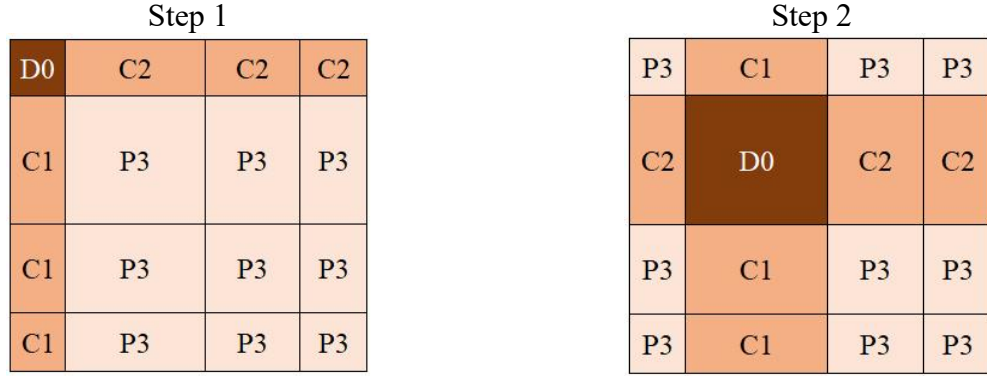


Figure 2. Illustration of *BFWUS* operation: cross moves from top-left to bottom-right corner of blocked matrix *B*; first, block *D0* is calculated through itself; second, blocks *C1* and *C2* are calculated through *D0*; third, blocks *P3* are calculated through *C1* and *C2*

Algorithm 2: Calculation of unequally sized blocks (*USBC*)

Input: A vector *S* of sizes of graph vertex subsets
Input: A blocked matrix $B[M \times M]$ of path distances
Input: A blocked matrix $P[M \times M]$ of previous vertices in shortest paths
Input: Indices *v*, *m* and *u* of vertex subsets
Output: Recalculated block $B_{v,u}$ of matrix *B*
Output: Recalculated block $P_{v,u}$ of matrix *P*

```

for  $k \leftarrow 1$  to  $S_m$  do
  for  $i \leftarrow 1$  to  $S_v$  do
    for  $j \leftarrow 1$  to  $S_u$  do
       $sum \leftarrow B_{v,m}(i, k) + B_{m,u}(k, j)$ 
      if  $B_{v,u}(i, j) > sum$  then
         $B_{v,u}(i, j) \leftarrow sum$ 
         $P_{v,u}(i, j) \leftarrow P_{m,u}(k, j)$ 
return  $B, P$ 

```

Computation of shortest paths between vertices within two clusters through neighbor cluster and interconnecting edges. In case of matrix $B[2 \times 2]$, the *BFWUS* algorithm calculates two diagonal and two non-diagonal blocks in the following way:

$$B_{11}[S_1 \times S_1] \leftarrow B_{11}[S_1 \times S_1] \otimes B_{11}[S_1 \times S_1] \quad (1)$$

$$B_{21}[S_2 \times S_1] \leftarrow B_{21}[S_2 \times S_1] \otimes B_{11}[S_1 \times S_1] \quad (2)$$

$$B_{12}[S_1 \times S_2] \leftarrow B_{11}[S_1 \times S_1] \otimes B_{12}[S_1 \times S_2] \quad (3)$$

$$B_{22}[S_2 \times S_2] \leftarrow B_{21}[S_2 \times S_1] \otimes B_{12}[S_1 \times S_2] \quad (4)$$

$$B_{22}[S_2 \times S_2] \leftarrow B_{22}[S_2 \times S_2] \otimes B_{22}[S_2 \times S_2] \quad (5)$$

$$B_{12}[S_1 \times S_2] \leftarrow B_{12}[S_1 \times S_2] \otimes B_{22}[S_2 \times S_2] \quad (6)$$

$$B_{21}[S_2 \times S_1] \leftarrow B_{22}[S_2 \times S_2] \otimes B_{21}[S_2 \times S_1] \quad (7)$$

$$B_{11}[S_1 \times S_1] \leftarrow B_{12}[S_1 \times S_2] \otimes B_{21}[S_2 \times S_1]. \quad (8)$$

Equations (1) – (4) calculate block B_{22} through block B_{11} . First, diagonal block B_{11} is calculated through itself using (1). Operator \otimes denotes a matrix MIN-PLUS multiplication operation. Then blocks B_{21} and B_{12} are calculated through B_{11} using (2) and (3). After that, diagonal block B_{22} is calculated through blocks B_{21} and B_{12} using (4). Equations (5) – (8) perform similar operations in opposite direction, i.e. from block B_{22} to block B_{11} . It is important that three operations are needed to calculate B_{22} through B_{11} , i.e. (2), (3) and (4). Similarly, three operations are needed to calculate B_{11} through B_{22} , i.e. (6), (7) and (8). To accomplish it, two intermediate blocks are additionally calculated, i.e. B_{21} and B_{12} . For large sparse graphs, *BFWUS* requires a huge amount of memory space and processor time. For very large graphs it is unjustified and practically unacceptable.

We propose a new method of computing B_{22} through B_{11} (and similarly B_{11} through B_{22}). The method allows to account for features of sparse graphs with clustered vertices, to reduce the amount of consumed memory space, to decrease the number of matrix MIN-PLUS operations executed over blocks.

Let two clusters $Clust_{11}$ and $Clust_{22}$ divide the vertex set V of graph G into two subsets V_1 and V_2 of unequal sizes S_1 and S_2 (Figure 3,a). The vertices of clusters $Clust_{11}$ and $Clust_{22}$ are connected by edges from Con_{12} and Con_{21} . As a result, we obtain a matrix B consisting of four blocks (Figure 3,b). Blocks B_{11} and B_{22} initially describe the internal weighted edges of the clusters, and then describe the shortest path lengths between vertices of set V_1 and between vertices of set V_2 . Sparse blocks W_{12} and W_{21} describe weighted edges connecting vertices of V_1 to vertices of V_2 and vice versa respectively. Figure 3,c shows that blocks B_{11} and B_{22} are placed in memory as matrices of distances (*DiM*) using row-major memory layout, and blocks W_{12} and W_{21} are placed as adjacent lists (*AjL*).

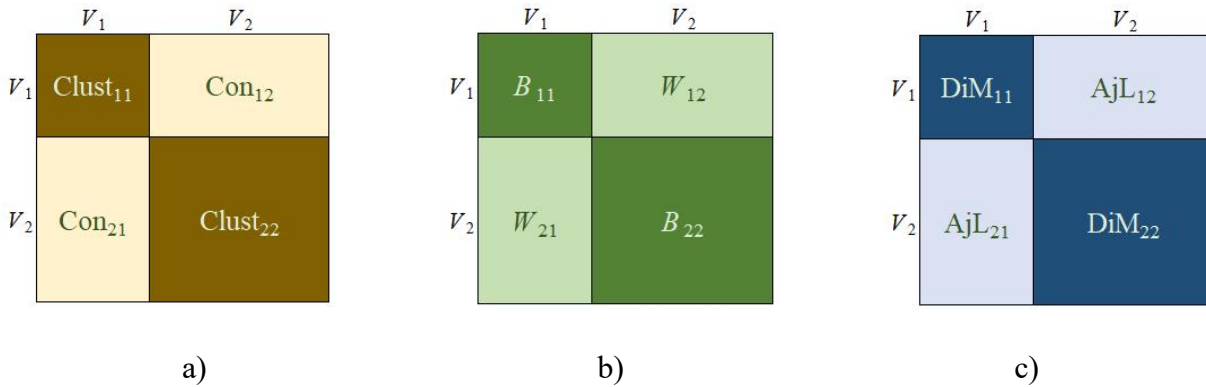


Figure 3. Dividing graph to clusters: a) diagonal blocks are clusters and non-diagonal blocks describe connections between clusters; b) blocks B_{11} and B_{22} describe shortest path lengths and blocks W_{12} and W_{21} describe weighted edges; c) diagonal blocks are represented by matrix of distances (*DiM*) and non-diagonal blocks are represented by adjacent lists (*AjL*)

Our method of computing block B_{22} through block B_{11} and vice versa consists in performing the following five operations:

$$\begin{aligned}
 B_{11} &\leftarrow \text{Diagonal}(B_{11}) & (9) \\
 B_{22} &\leftarrow \text{BlockThroughBlock}(B_{22}, W_{21}, B_{11}, W_{12}) & (10) \\
 B_{22} &\leftarrow \text{Diagonal}(B_{22}) & (11) \\
 B_{11} &\leftarrow \text{BlockThroughBlock}(B_{11}, W_{12}, B_{22}, W_{21}) & (12) \\
 B_{11} &\leftarrow \text{Diagonal}(B_{11}) & (13)
 \end{aligned}$$

Operation *Diagonal* (B_{ii}) calculates the shortest paths between all vertices of block B_{ii} ; the paths can traverse through edges within B_{ii} and outside it. The operation can be performed using (1) or can preferably be implemented by the fast *GEA* algorithm proposed in [16, 17].

New operation $BlockThroughBlock(B_{ij}, W_{ji}, B_{ii}, W_{ij})$ calculates the shortest paths between vertices of block B_{ij} . It traverses through edges of block W_{ji} , then through vertices and edges of block B_{ii} and finally through edges of block W_{ij} . Since the edges of blocks W_{ji} and W_{ij} are not numerous in the sparse graph, the $BlockThroughBlock$ operation is fast. When W_{ji} or W_{ij} is empty, $BlockThroughBlock$ is not executed at all. Moreover, the shortest paths of these blocks do not need to be stored in memory, only edge descriptions are needed to store. This is a big advantage of our method which yields accurate solutions. The shortest paths between vertices of different clusters are calculated in real time. It is easier to compute the shortest paths between clusters after computing shortest paths within clusters.

Approximate fast algorithm of computing all-pairs shortest paths in clusters of sparse graph. We have developed an approximate APSP algorithm operating on graph clusters (AAPSPC) aiming at reduction of the consumed memory space and CPU time. Figure 4,a shows the content of matrix B and Figure 4,b depicts the matrix representation and placement in memory. Clusters are placed in the matrix principal diagonal, and interconnections of clusters are represented by graph edges (blocks of matrix W) allocated outside of principal diagonal. Clusters are represented by matrices of shortest path distances (DiM). Blocks of matrix W are represented as adjacent lists (AjL).

B_{11}	W_{12}	W_{13}	W_{14}
W_{21}	B_{22}	W_{23}	W_{24}
W_{31}	W_{32}	B_{33}	W_{34}
W_{41}	W_{42}	W_{43}	B_{44}

a)

DiM ₁	AjL ₁₂	AjL ₁₃	AjL ₁₄
AjL ₂₁	DiM ₂	AjL ₂₃	AjL ₂₄
AjL ₃₁	AjL ₃₂	DiM ₃	AjL ₃₄
AjL ₄₁	AjL ₄₂	AjL ₄₃	DiM ₄

b)

Figure 4. Matrix B of shortest paths distances in clusters: a) diagonal blocks B_{ii} are clusters and non-diagonal blocks W_{ij} represent edges connecting clusters; b) each cluster is represented by distances matrix (DiM) and each non-diagonal block is represented by adjacent list (AjL)

Algorithm 4 specifies AAPSPC. Its input data are the vertex subsets corresponding to clusters, and their sizes. Matrices $B = W$ and P describing initial states of shortest paths in clusters are also input data. The algorithm output data are the completely calculated diagonal blocks of matrices B and P .

The algorithm consists of two loop nests of depth two. The first nest traverses the diagonal blocks from the left-top to right bottom corner of the matrices and calculates the shortest paths of each succeeding block through previous blocks. The second nest does the same work in reverse order. Totally M diagonal blocks are calculated. The overall number of the diagonal block recalculations AAPSPC performs is

$$Rec(B) = M \cdot (M + 1). \quad (14)$$

It is almost M times smaller than the number M^2 of block recalculations BFWUS performs.

Function $Diagonal^*(S, B, P, d)$ is implemented using operation $Diagonal(B_{dd})$. Function $BlockThroughBlock^*(S, B, P, d, p)$ is implemented using operation $BlockThroughBlock(B_{dd}, W_{dp}, B_{pp}, W_{pd})$.

Algorithm 4: Computing all-pairs shortest paths in clusters of sparse graphs (*AAPSPC*)

Input: Subsets $V = (V_1, \dots, V_M)$ of set V of graph clustered vertices
Input: A vector $S = (S_1, \dots, S_M)$ of sizes of graph vertex subsets
Input: A blocked matrix $B[M \times M]$ representing graph
Input: A blocked matrix $P[M \times M]$ of previous vertices in shortest paths
Output: Recalculated matrix $B[M \times M]$
Output: Recalculated matrix $P[M \times M]$

```

for  $d \leftarrow 1$  to  $M$  do
     $Diagonal'(S, B, P, d)$ 
    if  $d > 1$  then
        for  $p \leftarrow 1$  to  $d - 1$  do
             $BlockThroughBlock^*(S, B, P, d, p)$ 
for  $d \leftarrow M$  down to  $1$  do
     $Diagonal'(S, B, P, d)$ 
    if  $d > 1$  then
        for  $p \leftarrow d - 1$  down to  $1$  do
             $BlockThroughBlock^*(S, B, P, p, d)$ 
return  $B, P$ 

```

The calculation of shortest paths between vertices of different clusters is carried out in real time through the earlier computing of shortest paths between vertices within clusters.

Let's estimate the volume of storage needed for allocation of matrix B in algorithms *BFWUS* and *AAPSPC*. *BFWUS* uses the *DiM* representation of all blocks, Therefore, $storage(BFWUS) = N^2$. The amount of storage consumed by *AAPSPC* can be estimated as

$$storage(AAPSPC) = storage(Clusters) + storage(Edges), \quad (15)$$

$$storage(Clusters) = \sum_{c=1}^M S_c^2 \quad (16)$$

where c is a cluster; M is the number of clusters; S_c is the number of vertices in cluster c ; $storage(Edges)$ is the size of all non-diagonal blocks W_{ij} describing interconnect edges. For sparse graphs, $storage(Edges)$ is not huge.

Let's consider the bounding case when all clusters have the same size $S_c = N / M$. Then the amount of storage they need for placement is

$$storage(Clusters) = \left(\frac{N}{M}\right)^2 \cdot M = N^2 / M \quad (17)$$

It is M times smaller than *BFWUS* needs. If the clusters are of unequal sizes, the gain is reduced. We conclude that *AAPSPC* consumes a much smaller amount of memory compared to *BFWUS* in case of sparse graphs. It is a very big advantage in the case of searching for the shortest paths in very large sparse graphs.

At the same time, *AAPSPC* has a drawback: it considers only shortest paths that go from one cluster to another (maybe iteratively) and return to the former cluster. It accounts for not all paths passing through several clusters before returning to the source. It is important to know what inaccuracies arise in calculating the shortest path distances due to these restrictions. We have done experiments with *AAPSPC* with respect to the inaccuracies.

For instance, we provide results obtained on a directed weighted graph consisting of 4 clusters, 128 vertices and 1502 edges with the average weight of 55.54. Figure 5 shows that the inaccuracies in computing the shortest path lengths are small and vary in the range 1.1778 % down to 0.0 % depending on the weights of edges that connect clusters. The inaccuracies disappear at the average weight of 23 which is 41.4 % compared to the average weight of edge within clusters. *AAPSPC* yields an accurate shortest path distances between all pairs of vertices in all clusters for all interconnect edge weights larger than 41.4 %.

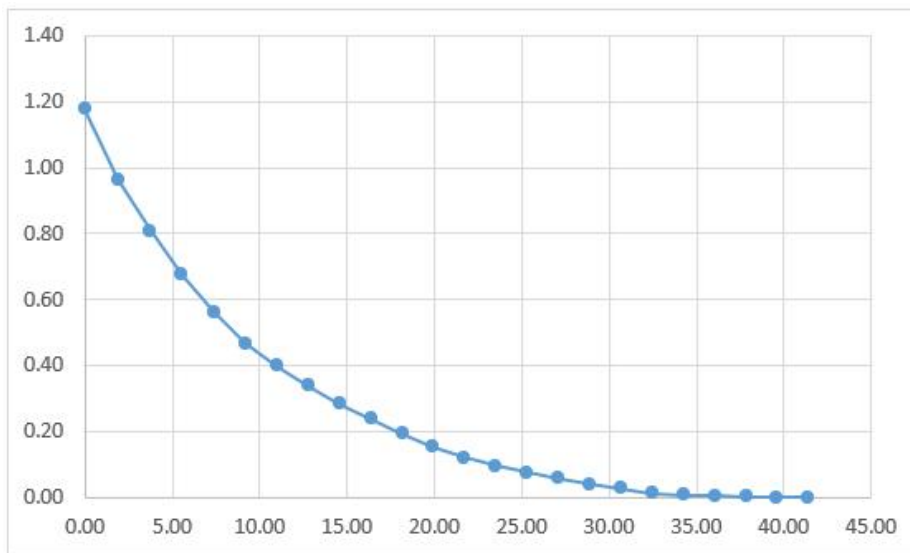


Figure 5. Inaccuracy (vertical axis, %) in computation of shortest path distances vs. share (horizontal axis, %) of average weight of edge between clusters in average weight of edge within clusters

Right reordering of clusters in matrix B can decrease the required number of block recalculations or / and reduce the inaccuracies in computations of shortest paths. Incorporating Dijkstra algorithm [1] or any algorithm of Dijkstra's family in *AAPSPC* is the way of avoiding the inaccuracies in shortest paths computation and obtaining accurate solutions.

Conclusion. The paper solves the all-pairs shortest paths problem on sparse graphs partitioned into clusters by finding dense weakly connected subgraphs. The approach we propose is based on recently published blocked algorithms which divide the graph into unequally sized subgraphs. We have developed an operation of computing the shortest paths between vertices of one cluster passing through vertices and edges of neighbor cluster, and through edges connecting the clusters. This allowed us to develop a fast and memory efficient approximate algorithm that first computes the shortest paths between vertices within each cluster and then computes shortest paths between vertices of different clusters in real time. The algorithm gives accurate solutions for road, computer and other networks in which the weights of edges connecting clusters are typically at least as large as the weights of edges within clusters.

References

- [1] Dijkstra E. W. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1959, vol. 1, no. 1, pp. 269–271.
- [2] Floyd R.W. Algorithm 97: Shortest path. *Communications of the ACM*, 1962, no. 5 (6), p. 345.
- [3] Glabowski M., Musznicki B., Nowak P. and Zwierzykowski P. Review and Performance Analysis of Shortest Path Problem Solving Algorithms. *International Journal on Advances in Software*, 2014, vol. 7, no. 1&2, pp. 20 – 30.
- [4] Madkour A., Aref W. G., Rehman F. U., Rahman M. A., Basalamah S. A Survey of Shortest-Path Algorithms. ArXiv: 1705.02044v1 [cs.DS], 4 May 2017, 26 p.

- [5] Prihozhy, A., Karasik, O. New blocked all-pairs shortest paths algorithms operating on blocks of unequal sizes. *System analysis and applied information science*. – 2023. – №. 4. – P. 4–13.
- [6] Rahman A.-H. Ab, Prihozhy A., Mattavelli M. Pipeline synthesis and optimization of FPGA-based video processing applications with CAL. *EURASIP Journal on Image and Video Processing*, vol. 2011:19, pp. 1–28.
- [7] Katz G. J., Kider J. T. All-pairs shortest-paths for large graphs on the GPU. *GH'08: Proceedings of the 23rd ACM SIGGRAPH/EUROGRAPHICS symposium on Graphics hardware*. ACM, 2008, pp. 47-55.
- [8] Ortega-Arranz H., Torres Y., Llanos D. R, and Escribano A. G. The all-pair shortest-path problem in shared-memory heterogeneous systems. *High-Performance Computing on Complex Environments*, 2013, pp. 283–299.
- [9] Djidjev H., Thulasidasan S., Chapuis G., Andonov R. and Lavenier D. Efficient multi-GPU computation of all-pairs shortest paths. *IEEE 28th International Parallel and Distributed Processing Symposium*. IEEE, 2014, pp. 360–369.
- [10] Venkataraman G., Sahni S., Mukhopadhyaya S. A Blocked All-Pairs Shortest Paths Algorithm. *Journal of Experimental Algorithmics (JEA)*, 2003, vol 8, pp. 857 – 874.
- [11] Park J. S., Penner M., and Prasanna V. K. Optimizing graph algorithms for improved cache performance. *IEEE Trans. on Parallel and Distributed Systems*, 2004, no. 15 (9), pp.769 – 782.
- [12] Bellman R. E. On a routing problem. *Quarterly of Applied Mathematics*, 1958, vol. 16, no. 1, pp. 87–90.
- [13] Johnson D. B. Efficient Algorithms for Shortest Paths in Sparse Networks. *J. ACM*, 1977, vol. 24 no. 1, pp. 1 – 13.
- [14] Harish P., Narayanan P. J. Accelerating large graph algorithms on the GPU using CUDA. *International conference on high-performance computing*. Springer, 2007, pp. 197–208.
- [15] Meyer U. and Sanders P. Δ -stepping: a parallelizable shortest path algorithm. *Journal of Algorithms*, vol. 49, no. 1, 2003, pp. 114–152.
- [16] Prihozhy A. A., Karasik O. N. Heterogeneous blocked all-pairs shortest paths algorithm. *System analysis and applied information science*, 2017, no. 3, pp. 68 – 75. (In Russian).
- [17] Prihozhy A.A., Karasik O.N. Inference of shortest path algorithms with spatial and temporal locality for big data processing. [*Big Data and Advanced Analytics: proceedings of VIII international conference*]. Minsk, Bestprint Publ., 2022, pp. 56 – 66.
- [18] Prihozhy A. A., Karasik O. N. Advanced heterogeneous block-parallel all-pairs shortest path algorithm. *Proceedings of BSTU, issue 3, Physics and Mathematics. Informatics*, 2023, no. 1 (266), pp. 77–83. DOI: 10.52065/2520-6141-2023-266-1-13.
- [19] Albalawi E., Thulasiraman P., Thulasiram R. Task Level Parallelization of All Pair Shortest Path Algorithm in OpenMP 3.0. *2nd International Conference on Advances in Computer Science and Engineering (CSE 2013)*. Los Angeles, CA, July 1 – 2, 2013, pp. 109 – 112.
- [20] Yang S., Liu X., Wang Y., He X., Tan G. Fast All-Pairs Shortest Paths Algorithm in Large Sparse Graph. *ICS '23: Proceedings of the 37th International Conference on Supercomputing*, 2023, pp. 277–288.
- [21] Prihozhy A. A. Simulation of direct mapped, k-way and fully associative cache on all-pairs shortest paths algorithms. *System analysis and applied information science*, 2019, no. 4, pp. 10 – 18.
- [22] Prihozhy A. A. Optimization of data allocation in hierarchical memory for blocked shortest paths algorithms. *System analysis and applied information science*, 2021, no. 3, pp. 40 – 50.
- [23] Prihozhy A.A., Karasik O.N. Cooperative block-parallel algorithms for task execution on multi-core system. *System analysis and applied information science*, 2015, no. 2, pp. 10–18.
- [24] Karasik O. N., Prihozhy A. A. Threaded block-parallel algorithm for finding the shortest paths on graph. *Doklady BGUIR*, 2018, no. 2, pp. 77 – 84. (In Russian).
- [25] Karasik O. N., Prihozhy A. A. Tuning block-parallel all-pairs shortest path algorithm for efficient multi-core implementation. *System analysis and applied information science*, 2022, no. 3, pp. 57 – 65.
- [26] Prihozhy A.A., Karasik O.N. Influence of shortest path algorithms on energy consumption of multi-core processors. *System analysis and applied information science*, 2023, no. 2, pp. 4-12.
- [27] Prihozhy A. A. Generation of shortest path search dataflow networks of actors for parallel multicore implementation. *Informatics*, 2023, vol. 20, no. 2, pp. 65–84.
- [28] Прихожий, А.А. Кооперативная модель оптимизации выполнения потоков на многоядерной системе / А.А. Прихожий, О.Н. Карасик // Системный анализ и прикладная информатика. – 2014. – № 4. – С. 13-20

author's contribution

Authors to make an equivalent contribution.

БЛОЧНЫЙ АЛГОРИТМ ПОИСКА КРАТЧАЙШИХ ПУТЕЙ В РАЗРЕЖЕННЫХ ГРАФАХ, РАЗБИТЫХ НА КЛАСТЕРЫ НЕРАВНОГО РАЗМЕРА

А.А. Прихожий

Профессор кафедры «Программное обеспечение информационных систем и технологий» Белорусского национального технического университета, д.т.н., профессор

О.Н. Карасик

Ведущий инженер иностранного производственного унитарного предприятия «ИССОФТ СОЛЮШЕНЗ» (ПВТ, г. Минск), к.т.н.

Аннотация. В данной статье рассматривается задача поиска кратчайших путей между всеми парами вершин ориентированного взвешенного разреженного графа, который разбит на кластеры путем нахождения плотных слабосвязных подграфов. Мы решаем эту задачу, разрабатывая новые блочные алгоритмы, которые описывают кратчайшие пути матрицами блоков неравных размеров, соответствующих размерам кластеров графа. Эти алгоритмы расширяют возможности известных существующих алгоритмов, использующих блоки одинакового размера (таких как блочный алгоритм Флойда-Уоршелла) в части адекватного моделирования графов реальных сетей различного назначения, а также в части эффективного использования параллелизма и вычислительных ресурсов многопроцессорных систем и многоядерных процессоров. Предлагаемый в данной статье блочный алгоритм поиска кратчайших путей в разреженных графах большого размера, разбитых на кластеры, позволяет, с одной стороны, сократить объем используемой памяти, а с другой – уменьшить количество пересчетов блоков. Диагональные блоки описывают кратчайшие пути внутри кластеров, недиагональные блоки описывают немногочисленные взвешенные дуги, соединяющие кластеры. Кратчайшие пути между вершинами разных кластеров вычисляются в реальном времени. Потребление памяти по сравнению с алгоритмами семейства Флойда-Уоршелла уменьшается в число раз, равное количеству кластеров. Чтобы уменьшить количество пересчетов блоков, нами предложена новая операция, позволяющая точно вычислить кратчайшие пути между вершинами одного кластера, проходящие через вершины и ребра другого кластера, а также через ребра, соединяющие кластеры. Применение только этой операции позволяет построить алгоритмы, которые находят решения, вносящие небольшую ошибку (несколько процентов) в длины кратчайших путей при малых весах ребер между кластерами, и позволяет находить точные решения при увеличении весов этих ребер. Точные решения могут быть получены для разреженных графов, моделирующих дорожные, компьютерные и другие сети.

Ключевые слова: разреженный граф, кластер, кратчайшие пути, блочный алгоритм, блоки неравного размера, разнородная система.

UDC 004.021:004.75

REQUIREMENTS TO METHODS OF GRAPH CLUSTERING AT THE AIM OF SOLVING THE SHORTEST PATH PROBLEM



O.N. Karasik

*Tech Lead at ISsoft Solutions (part of Coherent Solutions) in Minsk, Belarus,
PhD in Technical Science
karasik.oleg.nikolaevich@gmail.com*



A.A. Prihozhy

*Professor at the Computer and System Software Department,
Doctor of Technical Sciences,
Full Professor
Belarusian National Technical University
prihozhy@yahoo.com*

O.N. Karasik

Tech Lead at ISsoft Solutions (part of Coherent Solutions) in Minsk, Belarus; PhD in Technical Science (2019). Interested in parallel computing on multi-core and multi-processor systems.

A.A. Prihozhy

Full professor at the Computer and system software department of Belarusian national technical university, doctor of science (1999) and full professor (2001). His research interests include programming and hardware description languages, parallelizing compilers, and computer aided design techniques and tools for software and hardware at logic, high and system levels, and for incompletely specified logical systems. He has over 300 publications in Eastern and Western Europe, USA and Canada. Such worldwide publishers as IEEE, Springer, Kluwer Academic Publishers, World Scientific and others have published his works.

Abstract. *In this paper we considered utilization of graph clustering results in scope of solving all-pairs shortest path problem by means of blocked all-pairs shortest paths algorithm with unequally sized blocks. We defined a set of requirements for the results of graph clustering based on the inner working of the blocked algorithm. We have done an analysis of two existing, well-known graph clustering algorithms (Walktrap and Spinglass) to verify if existing clustering algorithm can produce results consumable by blocked all-pairs shortest path algorithm with unequally sized blocks. Our experiments show, that both algorithms can be used to produce compatible results, however, in different contexts.*

Keywords: *All-pairs shortest path problem, blocked algorithms, graph clustering, unequally sized blocks.*

Introduction. The problem of finding the shortest paths between vertices in a directed, weighted graphs, has numerous applications in many real-life domains, like traffics, computer networks, social networks, computer games, hardware benchmarking and so on [1–3]. The shortest paths problem can be formulate differently depending of what shortest paths have to be found: a problem of finding a shortest path between a pair of vertices (**Single Pair Shortest Path**), a problem of finding all shortest paths between one vertex and the rest of the vertices (**Single Source Shortest Path** or **Single Sink Shortest Path**) and a problem of finding all pairs of shortest paths between all pairs of vertices in the graph (**All Pairs Shortest Path**). Computational complexity of the shortest paths problem depends on multiple parameters such as size of the graph (number of vertices), density of the graph (number of edges), edge weight (real or integer,

positive or negative) and problem formulation. For each of the formulation and combination of parameters there are classical algorithms [4–7] developed.

However, effective solution of the problem on large graphs with classical algorithms might take a significant portion of time because they weren't designed to run on modern, multi-core, multi-processors, or heterogeneous systems. That is why, to reduce execution time modern algorithms (or modernized versions of the classic algorithms) try to make use of hardware parallelism on single chip, powerful accelerators, distributed systems and pre-computed ahead of time information about or combination of those [8–10].

In our previous works we concentrated on improving solution of all-pairs shortest path problem using blocked version of Floyd-Warshall algorithm [11] by exploiting parallelism, CPU caching, data dependencies and novel algorithms to recalculate blocks [12–15]. In this paper we are focused on defining requirements to use pre-computed information about target graph's clusters (clustering) to speed up execution.

Algorithms. In this section we provide a concise description of Floyd-Warshall [7] and blocked Floyd-Warshall [11] algorithms to lay the foundation for further analysis.

Floyd-Warshall algorithm operates on a cost adjacency matrix $D[N \times N]$, where N is the number of vertices in a graph, and element $D_{i,j}$ contains a weight of the edge between vertices i and j . The algorithm recalculates matrix D in N iterations, where every iteration consists of picking the vertex k (essentially all vertex of matrix D are picked) and checking if any of the paths between vertices i and j can be shortened through vertex k (see Figure 1).

```
int N = ...
...
function algorithm(matrix D)
  for k = 0 to N do
    for i = 0 to N do
      for j = 0 to N do
        D[i,j] = min(D[i,j], D[i,k] + D[k,j])
      end
    end
  end
end function
```

Figure 1. Pseudocode of original Floyd-Warshall algorithm

Blocked Floyd-Warshall algorithm extends Floyd-Warshall algorithm by splitting the matrix D into $S \times S$ blocks, effectively creating a matrix $B[M \times M]$, where $M * S = N$. The algorithm recalculates matrix B in M iterations, where every iteration consists of three phases: calculating the a single «diagonal» block, which has dependency on itself, $2 \cdot (M - 1)$ «cross» blocks, which have dependency on themselves and the “diagonal” block and $(M - 1)^2$ «peripheral» blocks, which have dependency on the corresponding vertical and horizontal «cross» blocks (see Figure 2).

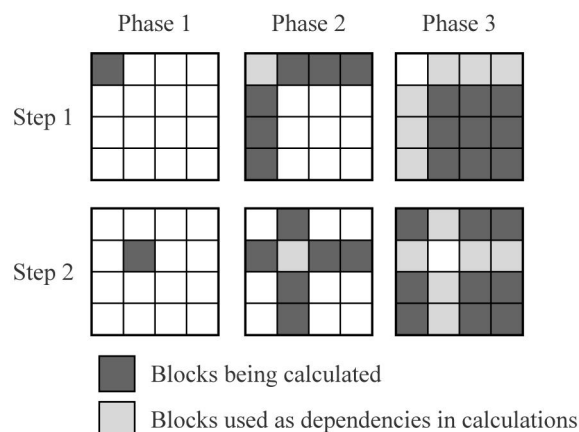


Figure 2. Illustration of calculation phases of Blocked Floyd-Warshall algorithm on first two iterations (steps)

All the blocks are calculated using the same procedure which always accepts three blocks (see Figure 3).

```

int M = ...
int S = ...
...
function blocked_algorithm(block_matrix B)
  for m = 0 to M do
    proc(B[m,m], B[m,m], B[m,m]);
    for i = 0 to m - 1 do
      proc(B[i,m], B[i,m], B[m,m]); proc(B[m,i], B[m,m], B[m,i]);
    end
    for i = m + 1 to M - 1 do
      proc(B[i,m], B[i,m], B[m,m]); proc(B[m,i], B[m,m], B[m,i]);
    end
    for i = 0 to m - 1 do
      for j = 0 to m - 1 do proc(B[i,j], B[i,m], B[m,j]); end
      for j = m + 1 to M - 1 do proc(B[i,j], B[i,m], B[m,j]); end
    end
    for i = m + 1 to M - 1 do
      for j = 0 to m - 1 do proc(B[i,j], B[i,m], B[m,j]); end
      for j = m + 1 to M - 1 do proc(B[i,j], B[i,m], B[m,j]); end
    end
  end
end function

function proc(B1, B2, B3)
  for k = 0 to S do
    for i = 0 to S do
      for j = 0 to S do
        B1[i,j] = min(B1[i,j], B2[i,k] + B3[k,j])
      end
    end
  end
end function

```

Figure 3. Pseudocodes of blocked Floyd-Warshall algorithm (blocked_algorithm) and block calculation procedure (proc).

The blocked version of the algorithms takes advantage of the CPU caching by improving data localization during block recalculations i.e. no more than three blocks participate in calculation, which makes it possible to fit all of them into CPU cache and reduce number of main memory accesses.

Clustering. In this section we provide a concise description of the graph clustering to lay a foundation for further analysis.

Graph clustering is an operation of grouping vertices of the graph into clusters (or communities) by taking into consideration the structure of the graph (edge adjacency, distance or

similarity between vertices, type of the graph and so on) in such a way that there should be many edges within each cluster and few between the clusters [16]. In terms of directed graphs, which are the main audience of the shortest path problem, graph clustering provides information about highly inter-connected and sparsely-connected sub-graphs, bridge vertices (i.e. vertices which constitute edges between clusters) and bridge edges (i.e. edges between clusters).

Depending on the graph, clustering algorithm and its parameters (or in cases when clustering algorithm uses any kind of randomization, even a particular run) the clustering results can have significant difference in number of clusters, number of vertices in clusters and number of bridge vertices and edges. In Table 1 you can find clustering results of the small (80 vertex) randomly generated, directed, connected graph using Walktrap [17] and Spinglass [18] clustering algorithms and on the Figure 4 you can see a visualization of execution of Walktrap, with *walks* set to 7, and Spinglass, with *spins* set to 7, algorithms.

Table 1. Results of the clustering of a randomly generated, directed, connected graph of 80 vertices and 94 edges using Walktrap and Spinglass clustering algorithms with different sets of parameters.

Walktrap				Spinglass			
Walks	Clusters	Bridge vertices	Bridge edges	Spins	Clusters	Bridge vertices	Bridge edges
2	17	24	26	2	2	7	7
3	19	26	28	3	3	10	10
4	19	27	29	4	4	11	11
5	19	26	29	5	5	13	13
6	18	27	28	6	6	15	15
7	20	29	30	7	7	17	17

Table 1 demonstrates how results of graph clustering differ in number of clusters and bridge vertices / edges between different algorithms and their parameters.

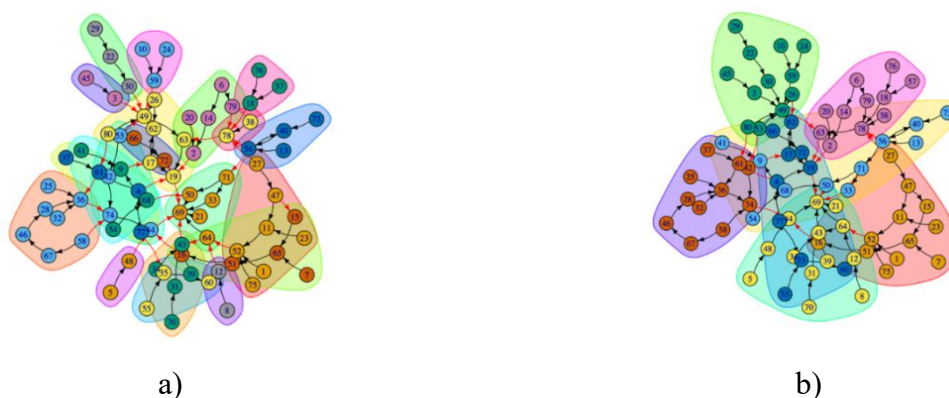


Figure 4. Visualization of graph clustering of a randomly generated, directed, connected graph of 80 vertices using Walktrap, with *walks* set to 7, (a) and Spinglass, with *spins* set to 7, (b) algorithms. Red arrows represent bridge edges, filled shapes and vertex colors represent clusters.

Analysis. In this section we perform an analysis of the results of execution of Walktrap and Spinglass graph clustering algorithms to define requirements to methods of graph clustering and input graphs based on how these results can be used to improve execution of blocked all-pairs shortest paths algorithm.

In appliance to blocked all-pairs shortest path algorithms the information from graph clustering can potentially be used to improve algorithm’s performance. However, because of how blocks are recalculated (see *Figure 2*) these optimizations (which might rely on advanced knowledge of the graph’s layout) are possible if clusters are matched to blocks i.e. every cluster is represented by a “diagonal” block. In classic variation of the blocked all-pairs shortest paths algorithm (see *Figure 3*), all the blocks must be of same size. This is a very strict requirement, which can hardly be matched by any of graph clustering algorithms. Luckily, this requirement can be relaxed as demonstrated in [19] by using blocks of unequal sizes. With an ability to use blocks of unequal sizes. Matching of clusters to “diagonal” blocks requires a rearrangement of rows and columns of cost adjacency matrix (before splitting it to blocks) to align all vertices of clusters around matrix diagonal. However, for the matching be successful and do not introduce a degradation instead of optimization, clusters must satisfy the following requirements:

1 *Clusters must not be too «small» or too «large»*⁴. Having a significant number of «small» clusters will result in a significant number of small blocks and CPU cache underutilization. Having large block results in contrary will reduce effect of the CPU cache and will result in a significant main memory pressure (as demonstrated in [20]).

2 *Clusters must not be too interconnected*. Having too many bridge vertices or bridge edges renders the makes the whole purpose of using clusters useless.

To understand, if the above requirements can be fulfilled, by existing graph clustering algorithm, we performed a set of experiments using two, well-known algorithms (Walktrap and Spinglass) on three randomly generated, connected graphs with artificially generated clusters (RCGWC). Information about all graphs is presented in Table 2. To verify how algorithms treat artificially generated clusters, during the generation of the graphs, we have recorded the ranges of vertices constituting to all clusters (see Table 3).

Table 2. Configuration of the experimental graphs, where RCG are randomly generated, connected graphs of different density and RCGWC are randomly generated connected graphs with artificially generated clusters.

Graph	Vertices	Edges	Bridge Vertices	Bridge Edges	Clusters
RCGWC-1	1200	49192	223	298	12
RCGWC-2	1200	46923	408	572	10
RCGWC-3	1200	56567	743	1415	10

We have executed all algorithms on all the experimental graphs with parameters varying from 2 to 12 walks and 2 to 12 spins for Walktrap and Spinglass algorithms respectively.

During the experiments, Walktrap algorithm has precisely identified all artificially generated clusters in all experimental graphs independently of number of walks. This demonstrates that results of graph clustering using Walktrap algorithm are very sensitive to graph’s layout and can be utilized when the layout matches the requirements. The executing time of the Walktrap algorithm varied from 0.3 to 1.5 seconds⁵ (depending on the number of walks – more walks more time).

⁴ The «small» and «large» here are relative to the characteristics of the executing CPU cache memory, where small is less than ¼ of Level 1 cache and “large” is more than ½ of Level 2 cache.

⁵ All clustering algorithms were executed on M1 Ultra Mac Studio using R language (igraph library).

Table 3. Configuration of artificially generated clusters of all RCGWC graphs.

Cluster	RCGWC-1		RCGWC-2		RCGWC-3	
	Vertices	Vertices Range	Vertices	Vertices Range	Vertices	Vertices Range
1	77	0 – 76	122	0 – 121	116	0 – 115
2	170	77 – 246	154	122 – 275	134	116 – 249
3	84	247 – 330	111	276 – 386	58	250 – 307
4	55	331 – 385	166	387 – 552	61	308 – 368
5	64	386 – 449	146	553 – 698	167	369 – 535
6	143	450 – 592	161	699 – 859	56	536 – 591
7	53	593 – 645	115	860 – 974	155	592 – 746
8	143	646 – 788	49	975 – 1023	141	747 – 887
9	95	789 – 883	96	1024 – 1119	162	888 – 1049
10	109	884 – 992	80	1120 – 1199	150	1050 – 1199
11	120	993 – 1112				
12	87	1113 – 1199				

During experiments (see Table 4), Spinglass algorithm produced various results. In most cases, the number of clusters was equal to number of spins (which is expected because of how algorithm works). However, when algorithm was executed against RCGWC-2 and RCGWC-3 graphs, which have 10 artificial clusters each, with number of spins set to 11 and 12 it still identified 10 clusters (not 11 or 12). Produces clusters only partially matched the structure of artificially generated clusters but despise this they weren't highly interconnected or of significantly different sizes (see Table 5):

- 1 The number of bridge vertices varied:
 - from 210 to 299 (where predefined was 223) for RCGWC-1
 - from 270 to 425 (where predefined was 408) for RCGWC-2
 - from 544 to 750 (where predefined was 743) for RCGWC-3
- 2 The number of bridge edges varied:
 - from 327 to 496 (where predefined was 298) for RCGWC-1
 - from 323 to 711 (where predefined was 572) for RCGWC-2
 - from 889 to 1522 (where predefined was 1415) for RCGWC-3

In combination, results presented in Table 4 and Table 5 experimentally demonstrate that results of graph clustering using Spinglass algorithm can be utilized with other parameters to optimally split the cost-adjacency matrix into block even when the natural layout of the graph doesn't matches the requirements. The executing time of the Spinglass algorithm varied from 9 to 11 seconds (depending on the number of spins – more spins more time).

Table 4. Experimental results of the Spinglass algorithm executed on all RCGWC graphs depicting number of clusters, number of bridge vertices and edges in relation to spins parameter.

Spins	RCGWC-1			RCGWC-2			RCGWC-3		
	Clusters	Bridge Vertices	Bridge Edges	Clusters	Bridge Vertices	Bridge Edges	Clusters	Bridge Vertices	Bridge Edges
2	2	224	353	2	270	323	2	544	889
3	3	210	327	3	360	506	3	618	1074
4	4	229	381	4	376	575	4	691	1288
5	5	253	448	5	392	584	5	697	1324

End of table 4

Spins	RCGWC-1			RCGWC-2			RCGWC-3		
	Clusters	Bridge Vertices	Bridge Edges	Clusters	Bridge Vertices	Bridge Edges	Clusters	Bridge Vertices	Bridge Edges
6	6	222	328	6	391	617	6	714	1349
7	7	299	497	7	405	622	7	724	1361
8	8	273	475	8	398	619	8	739	1433
9	9	268	473	9	414	711	8	747	1522
10	10	250	360	10	417	662	10	750	1492
11	11	272	421	10	416	634	10	747	1514
12	12	274	452	10	425	673	10	749	1516

Table 5. Experimental results of Spinglass algorithm on RCGWC-1 graph depicting number of vertices in clusters in relation to spins parameter value.

Spins/Cluster	2	3	4	5	6	7	8	9	10	11	12
1	592	380	344	171	309	145	144	144	143	121	52
2	608	361	325	197	143	144	162	143	136	83	143
3		459	359	258	170	178	170	138	87	143	77
4			172	364	143	170	144	170	78	143	95
5				210	250	196	191	110	120	170	170
6					185	247	173	121	118	87	120
7						120	96	96	109	95	65
8							120	130	95	77	86
9								148	144	65	110
10									170	107	143
11										109	86
12											53

Conclusion. We have defined a set of requirements for methods of graph clustering and produced clustering results to be useful for blocked all-pairs shortest paths algorithms. We performed an experimental analysis of two, well-known graph clustering algorithms (Walktrap and Spinglass) to verify if produced results match the requirements and can be used in blocked all-pairs shortest paths with unequal block sizes, or this task requires a specialized clustering algorithm which must be found or developed.

The experiments demonstrated that Walktrap algorithm can be used when natural layout of the graph matches the requirements. Spinglass algorithm has demonstrated an ability to produce variable results depending on the parameters. This shows that algorithm can be used even when natural layout of the graph doesn't match the requirements. Through these experiments we have confirmed that existing methods of graph clustering can produce results compatible with defined requirements, which makes implementation of the potential optimizations based on the availability of clustering information a prospective direction for future research.

References

- [1] Comparison of HPC Architectures for Computing All-Pairs Shortest Paths. Intel Xeon Phi KNL vs NVIDIA Pascal / M. Costanzo [et al.]. – Springer, 2020. – P. 37–49.
- [2] Anu, P. Finding All-Pairs Shortest Path for a Large-Scale Transportation Network Using Parallel Floyd-Warshall and Parallel Dijkstra Algorithms / P. Anu, M. G. (Kumar) // Journal of Computing in Civil Engineering. – 2013. – Vol. 27, №. 3. – P. 263–273.

- [3] Schrijver, A. On the history of the shortest path problem / A. Schrijver // Documenta Mathematica. – 2012. – Vol. 17, №. 1. – P. 155–167.
- [4] E. W. Dijkstra. A note on two problems in connexion with graphs / E. W. Dijkstra // Numerische Mathematik. – 1959. – Vol. 1, №. 1. – P. 269–271.
- [5] Bellman, R. On a routing problem / R. Bellman // Quarterly of applied mathematics. – 1958. – Vol. 16, №. 1. – P. 87–90.
- [6] Johnson, D. B. Efficient algorithms for shortest paths in sparse networks / D. B. Johnson // Journal of the ACM (JACM). – 1977. – Vol. 24, №. 1. – P. 1–13.
- [7] Floyd, R. W. Algorithm 97: Shortest Path / R. W. Floyd // Communications of the ACM. – 1962. – Vol. 5, №. 6. – P. 345–.
- [8] Efficient multi-GPU computation of all-pairs shortest paths / H. Djidjev [et al.]. – IEEE, 2014. – P. 360–369.
- [9] The all-pair shortest-path problem in shared-memory heterogeneous systems / H. Ortega-Arranz [et al.] // High-Performance Computing on Complex Environments. – 2013. – P. 283–299.
- [10] A scalable parallelization of all-pairs shortest path algorithm for a high performance cluster environment / T. Srinivasan [et al.]. – IEEE, 2007. – P. 1–8.
- [11] Venkataraman, G. A Blocked All-Pairs Shortest Paths Algorithm / G. Venkataraman, S. Sahni, S. Mukhopadhyaya // Journal of Experimental Algorithmics (JEA). – 2003. – Vol. 8. – P. 857–874.
- [12] Karasik, O. Tuning block-parallel all-pairs shortest path algorithm for efficient multi-core implementation / O. Karasik, A. Prihozhy // System analysis and applied information science. – 2022. – №. 3. – P. 57–65.
- [13] Karasik, O. Profiling of energy consumption by algorithms of shortest paths search in large dense graphs / O. Karasik, A. Prihozhy // Big Data and Advanced Analytics: сб. материалов IX Междунар. науч.-практ. конф., Минск. – Минск: BSUIR, 2023. – P. 44–50.
- [14] Prihozhy, A. Influence of shortest path algorithms on energy consumption of multi-core processors / A. Prihozhy, K. Oleg // System analysis and applied information science. – 2023. – №. 2. – P. 4–12.
- [15] Карасик, О. Н. Кооперативный многопоточный планировщик и блочно-параллельные алгоритмы решения задач на многоядерных системах / О. Н. Карасик. – Белорусский государственный университет информатики и радиоэлектроники, 2019.
- [16] Schaeffer, S. E. Graph clustering / S. E. Schaeffer // Computer science review. – 2007. – Vol. 1, №. 1. – P. 27–64.
- [17] Pons, P. Computing communities in large networks using random walks / P. Pons, M. Latapy. – Springer, 2005. – P. 284–293.
- [18] Reichardt, J. Statistical mechanics of community detection / J. Reichardt, S. Bornholdt // Physical review E. – 2006. – Vol. 74, №. 1. – P. 016110.
- [19] Prihozhy, A. New blocked all-pairs shortest paths algorithms operating on blocks of unequal sizes / A. Prihozhy, O. Karasik // System analysis and applied information science. – 2023. – №. 4. – P. 4–13.
- [20] Karasik, O. Parallel blocked all-pair shortest path algorithm: block size effect on cache operation in multi-core system / O. Karasik, A. Prihozhy // BIG DATA and Advanced Analytics: сб. материалов VIII Междунар. науч.-практ. конф., Минск, 11-12 мая 2022. – Минск: Бестпринт, 2022. – P. 28–38.

author's contribution

Authors to make an equivalent contribution.

ТРЕБОВАНИЯ К МЕТОДАМ КЛАСТЕРИЗАЦИИ ГРАФОВ С ЦЕЛЬЮ РЕШЕНИЯ ЗАДАЧИ О КРАТЧАЙШИХ ПУТЯХ

О.Н. Карасик

*Ведущий инженер иностранного
производственного унитарного
предприятия «ИССОФТ СОЛЮШЕНЗ»
(ПВТ, г. Минск),
к.т.н.*

А.А. Прихожий

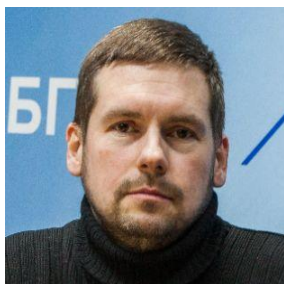
*Профессор кафедры «Программное
обеспечение информационных систем и
технологий» Белорусского
национального технического
университета, д.т.н., профессор*

Аннотация. В данной статье рассматривается возможность использования результатов кластеризации графа для решения задачи поиска всех кратчайших путей в графе при помощи блочного алгоритма поиска кратчайших путей, использующего блоки неравного размера. В статье определяются требования к результатам кластеризации графа на основании принципа работы блочного алгоритма поиска кратчайших путей. Проводится исследование двух, широко-известных алгоритмов кластеризации графа – *Walktrap* и *Spinglass*, с целью выяснения возможности использования результатов их работы блочным алгоритмом. Экспериментальные исследования показывают, что выбранные алгоритмы способны произвести результаты совместимые с определенными в статье требованиями, однако совместимость этих результатов во многом зависит от исходного графа и заданных параметров алгоритмов.

Ключевые слова: Кластеризация графа, блочные алгоритмы, поиск кратчайших путей, блоки неравного размера.

УДК [004.65:378]:17

ЭТИЧЕСКИЕ И СОЦИАЛЬНЫЕ АСПЕКТЫ ИСПОЛЬЗОВАНИЯ BIG DATA В СОВРЕМЕННОМ ОБРАЗОВАНИИ



П.А. Корбит
Магистр экономических наук, ассистент кафедры экономической информатики БГУИР
korbit@bsuir.by



Я.Е. Шепелевский
Студент кафедры инженерной психологии и эргономики, факультет компьютерного проектирования, БГУИР
yshepel1985@gmail.com

П.А. Корбит

Окончил Белорусский государственный университет информатики и радиоэлектроники (2001г.) и магистратуру Белорусского государственного университета информатики и радиоэлектроники (2020 г.) Защитил магистерскую диссертацию на тему «Цифровая трансформация высшего образования».

Я.Е. Шепелевский

Является студентом Белорусского государственного университета информатики и радиоэлектроники.

Аннотация. В статье рассматриваются этические и социальные аспекты применения больших данных в образовании, обсуждаются потенциал и проблемы использования данных в образовательном процессе. Автор анализирует, как большие данные могут улучшить качество образования, определяя уязвимости и адаптируя образовательный опыт. Также представлены методики и инструменты для анализа больших данных, а также проблемы, с которыми сталкиваются организации при их использовании, включая вопросы доступа, потери данных и конфиденциальности. Статья также затрагивает инновации в образовании через большие данные и развитие вычислительной педагогики как научно-инженерной дисциплины. Кроме того, обсуждается роль компьютерных систем в образовании и их возможности для экономии времени на поиск информации и преодоления образовательного неравенства.

Ключевые слова: большие данные, образование, этические аспекты, социальные проблемы, анализ данных, вычислительная педагогика, цифровые технологии, образовательное неравенство.

Введение. В современном мире цифровизация образования становится неотъемлемой частью образовательного процесса. Большие данные (*Big Data*) представляют собой уникальный ресурс, который может использоваться для улучшения качества образования, определения уязвимостей в образовательной системе и адаптации образовательного опыта для повышения эффективности. Однако, использование *Big Data* в образовании также поднимает ряд этических и социальных вопросов, которые необходимо учитывать при разработке и реализации соответствующих стратегий и политик.

Потенциал *Big Data* в образовании. Большие данные могут значительно повысить эффективность образовательного процесса, позволяя анализировать масштабные наборы данных для выявления тенденций и паттернов, которые могут быть использованы для

оптимизации учебных программ и методов. Например, анализ данных о поведении студентов может помочь в разработке персонализированных подходов к обучению, которые учитывают индивидуальные особенности каждого ученика.

Кроме того, *Big Data* может способствовать расширению доступа к образовательным возможностям. Дистанционное обучение и персонализированное образование становятся все более популярными благодаря возможности анализа больших объемов данных для выявления потребностей и предпочтений каждого студента.

Методики и инструменты анализа больших данных в образовании. Большие данные в образовании требуют специализированных методов анализа, чтобы извлечь ценную информацию и сделать ее полезной. Существует множество инструментов и методик, которые позволяют обрабатывать и анализировать большие объемы данных.

Data Science (наука о данных) играет ключевую роль в этом процессе, так как она объединяет различные техники и методы для обработки и интерпретации данных. Технологии машинного обучения и искусственного интеллекта также являются важными инструментами для анализа больших данных, так как они позволяют систематически обрабатывать и выявлять шаблоны и связи в данных, которые могут быть трудно обнаружить человеком [1].

Важным аспектом является классификация больших данных согласно теории «*VVV*», которая включает объем (*Volume*), скорость (*Velocity*) и разнообразие (*Variety*) обрабатываемых данных. Дополнительные категории включают «*4V*», добавляющее верифицируемость (*Veracity*), «*5V*» добавляющее жизнеспособность (*Viability*), и «*7V*», учитывающее изменчивость и визуализацию (*Variability and Visualization*) [1].

При работе с большими данными в образовании необходимо учитывать не только технические аспекты, но и этические и социальные вопросы. Это включает в себя вопросы конфиденциальности, доступа, а также возможные biases, которые могут возникнуть в результате анализа данных.

Среди современных методов анализа данных можно выделить статистический анализ, машинное обучение, анализ текста и анализ временных рядов. Эти инструменты позволяют обрабатывать большие объемы данных, выявлять сложные паттерны и делать точные прогнозы.

Применение этих методов в образовании может включать:

1 Анализ успеваемости студентов: Использование алгоритмов для анализа успеваемости позволяет выявить слабые места в учебном процессе и сфокусироваться на улучшении этих областей.

2 Персонализация обучения: С помощью анализа данных о поведении и предпочтениях студентов можно создавать индивидуальные учебные планы, которые учитывают уникальные потребности каждого ученика.

3 Создание инновационных образовательных продуктов: Использование данных для разработки новых методов обучения и образовательных материалов, которые могут быть более эффективными и адаптированы к современным требованиям.

Однако, при использовании больших данных в образовании необходимо учитывать и решать проблемы, связанные с их использованием. К таким проблемам относятся:

1 Конфиденциальность данных: При сборе и обработке больших данных о студентах необходимо обеспечивать их конфиденциальность и защищать личную информацию.

2 Недостаточное понимание данных: Для правильного анализа данных требуется определенная экспертиза в области статистики и анализа данных.

3 Неполнота данных: Большие данные могут быть неполными или содержать ошибки, что влияет на точность анализа.

4 Зависимость от технологий: Управление и анализ больших данных требуют современных технологий и инфраструктуры, что может быть дорогостоящим.

5 Недостаток человеческого фактора: Чрезмерное использование данных может привести к уменьшению роли преподавателей и снижению качества образования, если не учитывать индивидуальные особенности каждого студента.

Проблемы использования больших данных в образовании. Одна из ключевых проблем, с которой сталкиваются организации при использовании больших данных в образовании, – это доступ к данным. Необходимость в доступе к большому количеству информации может столкнуться с барьерами, такими как недостаток ресурсов для обработки данных или сложность интеграции разрозненных источников данных. Организациям нужно обеспечить четкую коммуникацию между различными участниками процесса, чтобы убедиться, что все имеют доступ к необходимым данным [2].

Существует риск потери данных, особенно если организации не вкладывают ресурсы в эффективные меры безопасности, такие как регулярные аудиты безопасности, многофакторная аутентификация, шифрование данных и управление патчами [2]. Это может привести к серьезным последствиям, включая ущерб репутации и финансовые потери.

Длительное хранение данных также может стать проблемой, поскольку организации должны соблюдать законодательство о защите данных, которое часто требует удаления или анонимизации данных после определенного периода времени. Это может потребовать дополнительных ресурсов для обеспечения соответствия законодательству [2].

Вопросы конфиденциальности являются центральными при работе с большими данными, так как неясные или недостаточно строгие декларации о конфиденциальности могут привести к нарушениям конфиденциальности. Чтобы предотвратить подобные ситуации, организации должны устанавливать четкие стандарты и процедуры для сбора, хранения и использования данных [2].

Чтобы успешно преодолеть эти проблемы, организации должны уделить внимание следующим решениям:

- установление и поддержание четких каналов коммуникации для всех участников процесса;
- проводить регулярные аудиты безопасности для выявления и устранения потенциальных уязвимостей;
- внедрение многофакторной аутентификации и шифрования данных для повышения безопасности;
- обеспечение регулярных резервных копий данных и их хранение в безопасных средах;
- ограничение доступа к данным только тем сотрудникам, которым он действительно необходим;
- избегание чрезмерной зависимости от инструментов конфиденциальности данных и поддержание человеческого суждений и понимания в процессе работы с данными [2].

Инновации в образовании через большие данные. Большие данные в образовании могут стать движущей силой инноваций, открывая новые горизонты для обучения и воспитания. Использование данных позволяет проводить глубокий анализ образовательных процессов и результатов, что в свою очередь ведет к созданию индивидуализированных образовательных программ и путей для каждого студента.

Одним из примеров инноваций, возможностей, которые открываются благодаря анализу больших данных, является метод портфолио. Этот метод подразумевает сбор и оценку работ студента в течение всего периода обучения, что позволяет визуализировать его прогресс и достижения, а также выявить области, где может потребоваться дополнительная поддержка.

Использование больших данных также способствует формированию индивидуальных образовательных маршрутов. С помощью анализа данных о предыдущих успехах и слабых сторонах студента можно создать персонализированный план обучения, который будет учитывать его уникальные потребности и интересы.

В целом, большие данные в образовании могут привести к новым открытиям, таким как определение факторов, влияющих на успеваемость учеников, выявление проблем в образовательной системе и разработка новых подходов к обучению. Это, в свою очередь, может повысить уровень удовлетворенности обучением, так как студенты будут видеть, что их усилия и индивидуальные особенности признаются и учитываются.

Инновации в образовании, поддерживаемые большими данными, могут также способствовать социализации школьников, сохранению психического и физического здоровья детей, а также обновлению образовательной организации в целом. Таким образом, большие данные в образовании являются неотъемлемой частью современного образовательного процесса, который стремится к гармоничному развитию подрастающего поколения [3-4].

Вычислительная педагогика и большие данные. Вычислительная педагогика – это направление, которое занимается изучением и применением компьютерных технологий в образовательном процессе. Оно акцентирует внимание на том, как компьютеры и компьютерные программы могут быть использованы для улучшения образовательного процесса и для создания новых подходов к обучению и воспитанию.

В рамках вычислительной педагогики большие данные играют ключевую роль. Использование больших данных позволяет анализировать и обрабатывать информацию о поведении и достижениях студентов, что в свою очередь может привести к новым открытиям в области образования. Например, анализ данных может выявить, какие методы обучения наиболее эффективны для разных групп студентов, и на основе этих данных можно адаптировать обучающие программы, чтобы они лучше соответствовали потребностям каждого отдельного ученика.

Одним из примеров применения вычислительной педагогики с использованием больших данных является система *Scratch*, которая использует визуальное программирование для обучения вычислительному мышлению. *Scratch* позволяет учащимся создавать свои цифровые истории, игры и модели, используя блоки кода, которые могут быть анализированы для оценки уровня формирования навыков вычислительного мышления. Это позволяет учителям и студентам лучше понять, какие навыки и стратегии используются в проектах, и как можно их улучшить.

Вычислительная педагогика также включает в себя развитие умений вычислительного участия и вычислительной рефлексии. Это означает, что ученики не только учатся использовать компьютеры и программы, но и учатся взаимодействовать с другими людьми и программами в цифровой среде, а также отражать и анализировать свои собственные действия и процессы обучения.

Вычислительная педагогика и большие данные вместе могут привести к созданию более интерактивных, персонализированных и эффективных образовательных программ. Однако, чтобы полностью реализовать потенциал этих технологий, необходимо учесть и решить ряд проблем, связанных с конфиденциальностью данных, доступом к информации и социальными аспектами использования больших данных в образовании.

Цифровые технологии и образование. Цифровые технологии играют центральную роль в современном образовании, предоставляя инструменты и ресурсы, которые значительно улучшают доступность и качество образовательных услуг. Они облегчают процесс обучения, позволяя студентам и преподавателям работать в гибких форматах и использовать различные платформы для улучшения взаимодействия и обучения.

Компьютерные системы и программное обеспечение в образовании могут помочь в экономии времени на поиск информации, предоставляя доступ к большим базам данных, онлайн-ресурсам и обучающим материалам. Это позволяет студентам и преподавателям сосредоточиться на активном обучении и развитии, вместо того чтобы тратить время на поиск и компиляцию информации.

Цифровые технологии также способствуют преодолению образовательного неравенства, обеспечивая доступ к образовательным возможностям для всех, независимо от их географического расположения или финансовых возможностей. Например, дистанционное обучение и онлайн-курсы позволяют учащимся из малообеспеченных районов и стран получить доступ к качественному образованию.

Однако, при внедрении цифровых технологий в образование важно учитывать вопросы конфиденциальности и безопасности данных, а также необходимость в обучении пользователей использованию новых инструментов и технологий. Возможность неправильного использования цифровых технологий может привести к нарушению конфиденциальности и безопасности данных, что может негативно сказаться на отношении к образовательным технологиям и их использованию в обществе.

В заключении можно сказать, что цифровые технологии представляют собой мощный инструмент для улучшения образования, но их успешное использование требует внимательного подхода к вопросам безопасности, доступа и обучения. При правильной реализации цифровые технологии могут значительно улучшить качество образования и сделать его более доступным для всех [5-6].

Заключение. В заключении можно подчеркнуть, что использование *Big Data* в образовании обладает большим потенциалом для улучшения качества образования и повышения уровня удовлетворенности обучением. Однако, важно учесть этические и социальные аспекты, связанные с использованием данных, чтобы обеспечить их безопасность и конфиденциальность, а также справедливость доступа к образовательным возможностям. Дальнейшие исследования в этой области будут способствовать развитию этих технологий в образовательном секторе, учитывая все вызовы и возможности, которые они представляют.

Список литературы

- [1] Меньшиков А, Перфильев В, Федосенко М, Фабзиев И. Основные проблемы использования больших данных в современных информационных системах М: Столыпинский вестник, 2022 – 329с.
- [2] Improvado – [Электронный ресурс]. – Режим доступа: <https://improvado.io/blog/big-data-analytics-privacy-problems> – дата доступа: 11.02.2024.
- [3] Молодой ученый – [Электронный ресурс]. – Режим доступа: <https://moluch.ru/archive/346/77841/> – дата доступа: 11.02.2024.
- [4] Nastroy – [Электронный ресурс]. – Режим доступа: <https://nastroy.net/post/sovremennyye-innovatsii-v-obrazovanii-vidyi-i-primeryi> – дата доступа: 11.02.2024.
- [5] Уваров А. Цифровая трансформация и сценарии развития общего образования. М.: Национальный исследовательский университет «Высшая школа экономики», 2020. – 109 с.
- [6] Субочева О, Яковлева Е. Цифровые технологии, как фактор формирования компетенций современного работника. М.: Журнал: «Общество: социология, психология, педагогика», 2020. – 4 с.

Авторский вклад

Корбит Павел Анатольевич – руководство исследованием по оценке этических и социальных аспектов использования Big Data в современном образовании.

Шепелевский Ярослав Евгеньевич – постановка задачи исследования, провел обширное исследование, включая анализ теоретической литературы, изучение примеров использования Big Data в образовательной практике и разработку концептуальных моделей, которые позволяют оценить эффективность и потенциал использования данных в образовании, разработка концепции вычислительной педагогики как научно-инженерной дисциплины, основанной на анализе больших данных, формирование структуры статьи.

ETHICAL AND SOCIAL ASPECTS OF USING BIG DATA IN MODERN EDUCATION

P.A. Korbit

*Master of Economics, Assistant of the Department of
Economic Informatics of BSUIR,*

Y.E. Shepelevsky

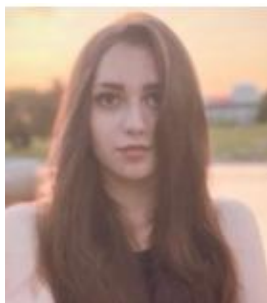
*Student of the Department of Engineering
Psychology and Ergonomics,
Faculty of Computer-Aided Design, BSUIR*

Abstract. The article examines the ethical and social aspects of the use of big data in education, discusses the potential and problems of using data in the educational process. The author analyzes how big data can improve the quality of education by identifying vulnerabilities and adapting educational experiences. Techniques and tools for analyzing big data are also presented, as well as the problems faced by organizations when using them, including issues of access, data loss and privacy. The article also touches on innovations in education through big data and the development of computational pedagogy as a scientific and engineering discipline. In addition, the role of computer systems in education and their possibilities for saving time on information retrieval and overcoming educational inequality are discussed.

Keywords: big data, education, ethical aspects, social issues, data analysis, computational pedagogy, digital technologies, educational inequality.

УДК 004.021:004.75

МЕТАВСЕЛЕННАЯ БЛОКЧЕЙНА



Е.И. Ковшер

Ассистент кафедры ЭВМ
БГУИР, магистрант
группы 356301
elizavetakovsher@mail.ru

Е.И. Ковшер

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием искусственного интеллекта и методов машинного обучения.

Аннотация. В данной статье представлен обзор концепции метавселенной блокчейна, рассмотрены ее технические аспекты, потенциал данной модели в решении актуальных проблем блокчейн-технологии. Помимо того, рассмотрены перспективы интеграции метавселенной блокчейна в различные отрасли, такие как финансы, здравоохранение и цифровые идентификации, что позволяет прояснить практическую значимость данного подхода в современном мире.

Ключевые слова: Блокчейн, метавселенная, интероперабельность, распределенные технологии.

Введение. В эпоху цифровой трансформации и инновационных технологий, блокчейн, как децентрализованная система хранения данных и обмена информацией, приобретает все большее значение в различных областях. Однако, несмотря на свою эффективность и прозрачность, традиционные блокчейн-системы сталкиваются с рядом ограничений, таких как масштабируемость и производительность [1].

В ответ на эти вызовы в научном сообществе появилась концепция «метавселенной блокчейна», представляющая собой инновационный подход к усовершенствованию блокчейн-технологии. Метавселенная блокчейна, или интероперабельная блокчейн-среда, стремится объединить несколько отдельных блокчейн-систем в единое пространство, где данные и активы могут свободно перемещаться и взаимодействовать.

Этот подход предлагает решение проблемы изоляции блокчейнов и создает новые возможности для совместного использования ресурсов. Вместо того чтобы ограничиваться отдельными цепями блоков, метавселенная блокчейна предоставляет расширенные возможности для управления данными, снижения нагрузки на сеть и повышения эффективности процессов [2].

В данной научной статье предпринята попытка глубже исследовать концепцию метавселенной блокчейна, рассмотреть ее технические аспекты, а также оценить потенциал данной модели в решении актуальных проблем блокчейн-технологии. Помимо того, рассмотрены перспективы интеграции метавселенной блокчейна в различные отрасли, такие как финансы, здравоохранение и цифровые идентификации, что позволяет прояснить практическую значимость данного подхода в современном мире.

Технология блокчейн. В последние годы блокчейн-технология стала центром внимания в мире информационных технологий, обещая революционные изменения в способе,

которым мы храним, передаем и подтверждаем данные. Основанный на децентрализации и прозрачности, блокчейн утверждается как эффективное решение многих проблем современных информационных систем [3].

Технология блокчейн представляет собой децентрализованный реестр или базу данных, обеспечивающую безопасный обмен информацией. В процессе осуществления транзакций, данные записываются в бухгалтерскую книгу, а обновление блокчейна осуществляется совместно с участием сети компьютеров. При возникновении новой транзакции она проходит проверку со стороны других компьютеров в сети, после чего непременно добавляется в реестр. Завершенный блок транзакций затем становится частью блокчейна, сохраняясь в нем в постоянной и неизменной форме.

Основное предназначение технологии блокчейн не ограничивается криптовалютами, она обладает множеством потенциальных применений. Одним из таких вариантов использования является управление цепочками поставок, где блокчейн обеспечивает высокую прозрачность перемещения товаров и услуг [4]. Технологию блокчейн также можно применять в системах голосования, где она гарантирует безопасность и прозрачность процесса голосования. Еще одним перспективным направлением использования является цифровая проверка личности, так как блокчейн способен создавать безопасные и защищенные от несанкционированного доступа цифровые удостоверения [5].

Возникновение концепции метавселенной блокчейна. Концепция метавселенной блокчейна возникла в контексте стремительного развития технологии блокчейн и криптовалют. Этот термин объединяет идею создания универсальной среды, объединяющей различные блокчейн-сети в единое цифровое пространство. Происхождение концепции метавселенной связано с несколькими ключевыми факторами.

Во-первых, с появлением множества различных блокчейн-платформ, каждая из которых предлагает свой уникальный протокол и функционал. Это создает фрагментацию в среде блокчейна, что затрудняет взаимодействие между разными сетями. Интероперабельность стала ключевой проблемой для дальнейшего развития технологии.

Во-вторых, стремление к созданию более масштабируемых и эффективных решений для существующих проблем в блокчейн, таких как масштабируемость, скорость транзакций и управление данными. Метавселенная блокчейна предлагает решение этих проблем путем создания универсального слоя, который может объединять различные блокчейн-сети, обеспечивая при этом высокую производительность и масштабируемость.

Третий фактор – это стремление к расширению возможностей смарт-контрактов и децентрализованных приложений (*DApps*). Метавселенная блокчейна предоставляет платформу, где смарт-контракты и *DApps* могут взаимодействовать с разными блокчейн-сетями, расширяя свои функциональные возможности и повышая их привлекательность для широкой аудитории.

Эволюция концепции метавселенной блокчейна происходит параллельно с разработкой технологий, таких как межцепочечные протоколы, которые позволяют взаимодействовать различным блокчейнам без потери безопасности и эффективности [6]. В результате рождается новый уровень децентрализации, открывая путь для создания более сложных и интегрированных экосистем, способных преодолевать текущие ограничения технологии блокчейн.

Технические аспекты метавселенной блокчейна. Технические аспекты метавселенной блокчейна обеспечивают функциональность и взаимодействие между различными блокчейн-сетями, что создает более универсальное и эффективное цифровое пространство. Эти аспекты включают в себя использование межцепочечных протоколов, стандартизированных протоколов смарт-контрактов, управление идентификацией, консенсус-протоколов и обеспечение масштабируемости [7].

Межцепочечные протоколы (*Interoperability Protocols*): одним из ключевых технических аспектов метавселенной блокчейна является использование межцепочечных протоколов. Эти протоколы действуют как мосты, обеспечивая связь и обмен данными между различными блокчейн-сетями. Такое взаимодействие позволяет пользователям и разработчикам проводить транзакции и передачу данных между разными цепями блоков, создавая единое цифровое пространство.

Стандартизированные протоколы смарт-контрактов: метавселенная блокчейна стремится к унификации протоколов смарт-контрактов для обеспечения их совместимости и легкости в разработке. Это позволяет децентрализованным приложениям, работающим на одной блокчейн-сети, быть эффективно развернутыми и функционировать на других цепях блоков без значительных изменений, способствуя более гибкой и расширяемой экосистеме.

Управление идентификацией и консенсус-протоколы: в метавселенной блокчейна управление идентификацией становится ключевым компонентом для обеспечения безопасности и подтверждения легитимности участников. Консенсус-протоколы, которые обеспечивают согласованность и достоверность данных между различными блокчейнами, играют также важную роль в обеспечении целостности метавселенной.

Сквозная безопасность (*Cross-Chain Security*): для защиты данных и транзакций, пересекающих различные блокчейн-сети, метавселенная блокчейна внедряет сквозные меры безопасности. Это включает в себя межцепочечные аудиты, защиту от атак и другие меры, гарантирующие безопасность взаимодействия между цепями блоков.

Масштабируемость (*Scalability*): обеспечение масштабируемости является еще одним важным техническим аспектом метавселенной блокчейна. С увеличением числа участников и транзакций необходимо обеспечивать высокую производительность и эффективное использование ресурсов. Разработка масштабируемых архитектур и оптимизация протоколов играют ключевую роль в поддержании стабильной работы метавселенной.

Метавселенная блокчейна в различных отраслях. Метавселенная блокчейна может быть полезна в различных отраслях, где требуется эффективное взаимодействие между разными блокчейн-сетями и расширение функциональности децентрализованных приложений. Некоторые из отраслей, где метавселенная блокчейна может иметь значительное воздействие, включают:

- финансовые услуги: обеспечение международных переводов и расчетов между различными финансовыми институтами.
- логистика и цепочка поставок: улучшение прозрачности и отслеживаемости товаров на каждом этапе цепочки поставок.
- здравоохранение: обеспечение безопасного обмена медицинской информацией между различными блокчейн-сетями.
- интернет вещей (*IoT*): создание универсального протокола для взаимодействия и обмена данными между устройствами *IoT* на различных блокчейн-платформах, обеспечение безопасности и конфиденциальности данных в сети *IoT*.
- энергетика: оптимизация и отслеживание производства энергии на различных блокчейн-сетях.
- образование: развитие децентрализованных образовательных приложений с возможностью взаимодействия в метавселенной.
- игровая индустрия: обеспечение обмена цифровыми активами и валютой между различными игровыми мирами и платформами.
- государственные услуги: внедрение универсальных систем идентификации граждан для обеспечения безопасного взаимодействия в электронных государственных сервисах [8, 9, 10].

Заключение. Сочетание метавселенной и технологии блокчейн предоставляет уникальные возможности для инноваций, открывая новые горизонты в различных сферах. Виртуальные торговые площадки становятся местами, где участники могут взаимодействовать, обмениваться цифровыми активами и создавать собственные виртуальные экосистемы. Игры, в свою очередь, приобретают новый уровень реализма и вовлеченности, предоставляя игрокам возможность зарабатывать реальные цифровые активы.

Однако с этими возможностями появляются и вызовы. Проблемы безопасности, конфиденциальности данных, а также вопросы, связанные с регулированием, могут стать барьерами на пути развития. Важно уделять внимание этическим аспектам использования данных в метавселенных и блокчейне, чтобы обеспечить безопасность и права участников.

Участие в активной разработке этих технологий позволяет не только следить за новыми тенденциями, но и активно влиять на их развитие. Совместные усилия сообщества, разработчиков и регуляторов могут помочь создать устойчивую и ответственную экосистему. Это важно для того, чтобы метавселенная и технология блокчейн могли стать средством, содействующим сотрудничеству и созиданию по всему миру.

Список литературы

- [1] Тапскотт Д. Технология блокчейн – то, что движет финансовой революцией сегодня / Тапскотт Дон. – М.: Эксмо, 2017 – 490 с.
- [2] Алабина Т.А., Дзангиева Х.С., Юшкова А.А. Метавселенная как глобальный тренд экономики // Экономика, профессия, бизнес. 2022. № 1 – С. 5-12.
- [3] Мелани С. Блокчейн. Схема новой экономики / Мелани Свон. – М.: Олимп-Бизнес, 2015 – 185 с.
- [4] Юхно А.С., Генкин А.С. Актуальные вопросы применения технологии блокчейн в контексте современного корпоративного управления. Управление риском. 2021. № 1 – С. 45-52.
- [5] Пестунов А.И. Криптовалюты и блокчейн: потенциальные применения в государстве и бизнесе // ЭКО. 2018. № 8 – С. 78-92.
- [6] Trunfio M., Rossi S. Advances in Metaverse Investigation: Streams of Research and Future Agenda // Virtual Worlds. 2022. № 1 – С. 103-129.
- [7] Dwivedi Y.K., Hughes L., Baabdullah A.M., Ribeiro-Navarrete S. Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy // International Journal of Information Management. 2022 – 66 с.
- [8] Салех К.С., Шарапова Н.В. Метавселенная как новая форма взаимодействия в виртуальной среде и ее возможности в экономическом развитии // Modern Economy Success. 2023. № 1 – С. 74-79.
- [9] Бауэр В.П. Блокчейн как основа формирования дополненной реальности в цифровой экономике // В.П. Бауэр, С.Н. Сильвестров, П.Ю. Барышников // Информационное общество. 2017. № 3 – С. 30-40.
- [10] Могайар У. Блокчейн для бизнеса / Уильям Могайар. – М.: Эксмо, 2016 – 382 с.

Авторский вклад

Ковшер Елизавета Игоревна – постановка задачи, обзор ключевых аспектов и перспектив развития метавселенной блокчейна, анализ полученных результатов.

BLOCKCHAIN METAVERSE

E. Kovsher

*Assistant at the Department of
Electronic Computing Machines,
JIR, Master's student group 356301*

Abstract. This article provides an overview of the concept of the blockchain metaverse, examining its technical aspects and exploring its potential in addressing current challenges within blockchain technology. Additionally, it delves into the prospects of integrating the blockchain metaverse into various industries, including finance, healthcare, and digital identities. This discussion helps elucidate the practical significance of this approach in the contemporary world.

Keywords: Blockchain, metaverse, interoperability, distributed technologies.

УДК 004.93'11, 004.032.26

АВТОМАТИЧЕСКАЯ ДИАГНОСТИКА СЕРДЕЧНОЙ АРИТМИИ ВО ВСТРОЕННЫХ СИСТЕМАХ



В.А. Реджепов

Аспирант кафедры электронных
вычислительных машин, БГУИР
vladimir.redzhepov@gmail.com



Д.Ю. Перцев

Доцент кафедры электронных
вычислительных машин, БГУИР,
кандидат технических наук,
доцент
pertsev@bsuir.by

В.А. Реджепов

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой методов и алгоритмов цифровой обработки сигналов и изображений.

Д.Ю. Перцев

Окончил аспирантуру Белорусского государственного университета информатики и радиоэлектроники (2016), защитил диссертацию в совете 05.13.01 (2020). Область научных интересов связана с разработкой методов и алгоритмов в системах технического зрения.

Аннотация. Техническое развитие носимых устройств привело к появлению простых и компактных встраиваемых систем мониторинга ЭКГ. При этом анализ результата является одним из ключевых показателей эффективности для коммерческих систем мониторинга. Это связано с тем, что классификация, как один из элементов анализа, позволяет идентифицировать аномальный сигнал ЭКГ и, следовательно, предложить соответствующее лечение. Поскольку для этого необходимы вычислительные мощности, обычно рассматривается клиент-серверный подход. Это означает, что встраиваемая система является лишь системой записи и хранения данных в традиционной системе мониторинга ЭКГ. Однако для обеспечения диагностики в режиме реального времени была введена необходимость классификации ЭКГ во встроженных системах. В данной статье рассматриваются существующие методы автоматической диагностики сердечных аритмий и предлагается решение для идентификации сердечной аритмии во встроженных системах.

Ключевые слова: ЭКГ, сердечная аритмия, цифровая обработка сигнала, фильтрация, реальное время, машинное обучение, классификация, *TinyML*, встроженные системы.

Введение. Сердечно-сосудистые заболевания (ССЗ) являются основной причиной смерти во всем мире, от которой по оценкам каждый год умирает 17,9 млн. человек. ССЗ представляют собой группу болезней сердца и кровеносных сосудов, в которую входят ишемическая болезнь сердца, заболевания сосудов головного мозга, ревматическая болезнь сердца и другие патологии. Более четырех из пяти смертей от ССЗ происходит в результате инфаркта и инсульта, причем треть из этих случаев смерти носит преждевременный характер и отмечается среди людей в возрасте до 70 лет [1]. В диагностике ССЗ важное место занимает мониторинг электрокардиограммы (ЭКГ) на наличие аритмий.

Сердечная аритмия – патологическое состояние, приводящее к нарушению частоты, ритмичности и последовательности возбуждения и сокращения сердечной мышцы –

миокарда. Аритмия – любой ритм сердца, отличающийся от нормального синусового ритма. При таком патологическом состоянии может существенно нарушаться нормальная сократительная активность сердца, что, в свою очередь, может привести к целому ряду серьёзных осложнений. Правильным, или регулярным, является синусовый ритм, который в пределах наблюдения задаётся только активностью синусового узла, правильный ритм которого принято называть “нормальным синусовым ритмом”, если он попадает в диапазон 60-90 ударов в минуту. Небольшие колебания периода сердечных сокращений, менее 0,1 секунды, считаются нормальной (физиологической) синусовой аритмией, связанной с естественной вариабельностью ритма сердца; их не считают патологическим нарушением. Отклонения от нормального состояния считаются нарушением, то есть аритмиями. Термин «аритмия» объединяет различные по механизму, клиническим проявлениям и прогностическому значению нарушения образования и проведения электрического импульса [2].

Для увеличения вероятности обнаружения аритмий необходимо проводить анализ суточной записи ЭКГ. Исследование представляет собой непрерывную регистрацию электрокардиограммы в течение 24 часов и более (48, 72 часа, иногда до 7 суток). Запись ЭКГ осуществляется при помощи специального портативного аппарата – рекордера (регистратора), который пациент носит с собой (на ремне через плечо или на поясе) [3]. Данный подход требует, чтобы устройство содержало достаточный объем памяти для хранения записей ЭКГ. Затем записанные данные передаются на сервер, где осуществляется их предварительная обработка и классификация автоматическими методами или медицинским персоналом. В таком случае пациент узнает о результатах исследования только после анализа записей [4]. Классификация, осуществляемая встроенной системой, позволяет получить предупреждение, что в сигнале ЭКГ есть аномалии в режиме реального времени. Портативные системы данного класса, как правило, не обладают большими объемами памяти и предполагают применение специальных методов, позволяющих использовать ограниченные ресурсы вычислительной системы. В последнее время стало возможно использовать технологии машинного обучения (ML) во встраиваемых системах на базе микроконтроллеров с низким энергопотреблением. Данный подход позволяет обучить модель на вычислительном кластере, а затем интегрировать ее во встроенную систему.

Классификация образов и диагностические решения. Конечной целью анализа биомедицинских сигналов является классификация данного сигнала с целью отнесения его к одной из нескольких известных категорий и получение диагностического решения, относящегося к состоянию пациента. Врач или медицинский специалист может достигнуть этой цели путём визуального или слухового анализа представленного сигнала: ему оказывает помощь сравнительный анализ данного сигнала с другими, для которых известны диагностические заключения. Кроме того, специалисты используют общепринятые протоколы и наборы правил. Существенную роль в этом процессе играют базовые знания, клинический опыт, квалификация и интуиция врача. Некоторые виды измерений, которые могут быть сделаны по данному сигналу, также могут оказать помощь в его анализе. Например, врач может измерить по графику ЭКГ ширину QRS-комплекса.

Когда анализ сигнала выполняется с использованием компьютерных алгоритмов, типичным результатом является получение некоторого набора численных параметров. Когда численные параметры относятся непосредственно к характеристикам сигнала, таким как ширина QRS-комплекса или длительность RR-интервала в сигнале ЭКГ, клинический специалист имеет возможность использовать эти параметры в логике принятия диагностических решений. Однако, когда вычисляются такие параметры, как коэффициенты авторегрессионной (AR) модели или спектральные статистики, аналитик вряд ли окажется в состоянии понять и использовать их. Более того, по мере увеличения числа рассчитываемых

ых компьютером параметров связанная с ними диагностическая логика становится слишком сложной и неприемлемой для анализа, выполняемого человеком. В таких случаях желательно, чтобы процессы классификации и принятия решений могли выполняться компьютером.

Прежде всего, необходимо иметь в виду, что биомедицинские сигналы являются лишь одним из источников информации для постановки диагноза: классификация данного сигнала с точки зрения его принадлежности к одной из многочисленных категорий может помочь в процедуре диагностики, но почти никогда не является единственным фактором. Тем не менее, классификация образов, основанная на анализе сигнала, действительно является важным аспектом анализа биомедицинских сигналов [5].

Базы данных для обучения алгоритмов. Согласно последним исследованиям [6], для отработки точности алгоритма, проверки и обучения систем искусственного интеллекта существует ряд международных открыто публикуемых баз данных электрокардиографических сигналов:

- *PTB Diagnostic ECG Database*;
- *MIT-BIH Arrhythmia Database*;
- *St Petersburg INCART 12-lead Arrhythmia Database*;
- *MIT-BIH Supraventricular Arrhythmia Database*;
- *AHA Database*;
- *UofT ECG database*;
- *Telemetric and Holter ECG Warehouse (THEW)*;
- *PHYSIOBC*.

PTB Diagnostic ECG Database состоит из 549 записей, полученных от 290 пациентов. Возраст составляет от 17 до 87 лет, что дает подробную информацию на уровне пациента, в том числе возраст, пол, диагноз, положительный медицинский анамнез, лекарства, предыдущие операции, наличие ишемической болезни сердца или любого другого сердечного заболевания. В этой базе данных используются ЭКГ с 12 отведениями и отведения Франка (x , y и z). Разрешение составляет 16 бит, 17 каналов (15 для ЭКГ, 1 для дыхания и 1 для линии напряжения), максимальное напряжение шума 10 мкВ [7].

MIT-BIH Arrhythmia Database находится в ведении Отделения медицинских наук и технологий Массачусетского технологического института в Центре биомедицинской инженерии. Он состоит из 48 получасовых записей, каждая из которых была получена от 47 пациентов в период с 1975 по 1979 год. Эти записи были оцифрованы с использованием 360 отсчетов в секунду на канал с разрешением 11 бит и диапазоном 10 мВ. Данные, содержащиеся в каждом долье пациента, включают дату регистрации, возраст, пол и назначенные лекарства [8].

St Petersburg INCART 12-lead Arrhythmia Database состоит из 75 записей электрокардиографии, извлеченных из 32 записей Холтера, которые представляют в общей сложности 32 пациента (17 мужчин и 15 женщин) в возрасте от 18 до 80 лет и средним возрастом 58 лет. Ни у одного из пациентов нет имплантированного кардиостимулятора, но у них есть различные проблемы с сердцем, включая ишемию миокарда, ишемическую болезнь сердца, нарушения проводимости и аритмии. Были использованы ЭКГ с 12 отведениями с частотой дискретизации 257 выборок в секунду и продолжительностью 30 минут [9].

MIT-BIH Supraventricular Arrhythmia Database была получена в больнице *MIT-Beth Israel Hospital (MIT-BIH)* и содержит 78 записей ЭКГ продолжительностью полчаса, выбранных в качестве примеров наджелудочковых аритмий. Записи были получены между 1990 и 1992 годами, с последующими аннотациями в 1999, 2010 и 2012 годах. В качестве методологии использовалась система *HOBBES (Hypothetical Obliterated Beat Sequences)* [10].

AHA Database доступна только на веб-сайте Научно-исследовательского института неотложной помощи (*ECRI Institute*) в виде запрошенного *DVD* вместе с оплатой пошлины. Включает 154 записи, разделенных на 8 типов аритмий: *Non-PVCs*, *Uniformly Isolated PVCs*, *Isolated Multiform PVCs*, *Bigeminus*, *R-T beats*, *Couplets*, желудочковые ритмы, фибрилляция желудочков или трепетание желудочков. Каждая запись *ЭКГ* (одна на пациента) длится в общей сложности 3 часа, разделенных на периоды не менее 30 минут, каждый для классификации сердечных сокращений. Информация была оцифрована в двух каналах *ЭКГ* с частотой 250 отсчетов в секунду и точностью 12 бит [11].

UofT ECG database состоит из выборки из 1020 пациентов, у которых *ЭКГ* в одном отведении была записана в течение 2-5 минут. Измерения проводились в различных позах тела (стоя, сидя, на треноге и лежа на спине) вместе с физическими упражнениями. Подгруппа из 43 пациентов посетила 6 сеансов записи *ЭКГ* в течение 6 месяцев. Отведения помещали на ладони и пальцы, получая отведение *DI* [12].

THEW – это организация, обслуживаемая Программой наблюдения за сердечно-сосудистыми исследованиями в Медицинском центре Университета Рочестера, проект начался в 2008 году. Он содержит более 3700 оцифрованных сигналов *ЭКГ*, в основном полученные через 24 часа холтеровского мониторинга. Эти записи *ЭКГ* являются частью 13 независимых исследований, которые доступны только участникам. В одном из исследований пациентов с острым инфарктом миокарда участвовало 90 пациентов со 160 записями, которые были взяты из 3 квазиортогональных отведений. Частота дискретизации составляет 200 отсчетов в секунду с амплитудой 10 мкВ [13].

База *PHYSIOBC* разработана в Мексике. Она состоит из записей для 91 пациента и всего содержит 182 записей; возраст пациентов был от 18 до 70 лет. Частота выборки для *ЭКГ* с 12 отведениями составляет 1000 выборок в секунду, а для *ЭКГ* Франка – 500 выборок в секунду. Продолжительность записи составляет от 2 до 5 минут на пациента [14].

Методы распознавания аритмий. Исторически распознавание аритмии проводилось стандартными методами логического и математического анализа цифровых сигналов *ЭКГ*. Однако с развитием машинного обучения задача по распознаванию определенной формы сигнала сердца может быть рассмотрена как классическая задача распознавания образов. Методы анализа аритмий можно разделить на три группы (рисунок 1) [15].

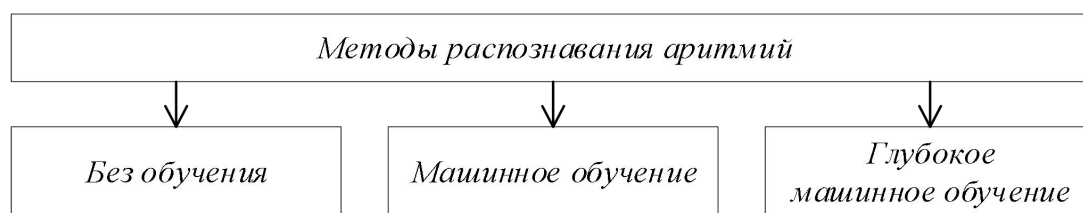


Рисунок 1. Классификация алгоритмов распознавания аритмий

В свою очередь, можно выделить основные этапы обработки данных алгоритмами из представленных групп (рисунок 2).

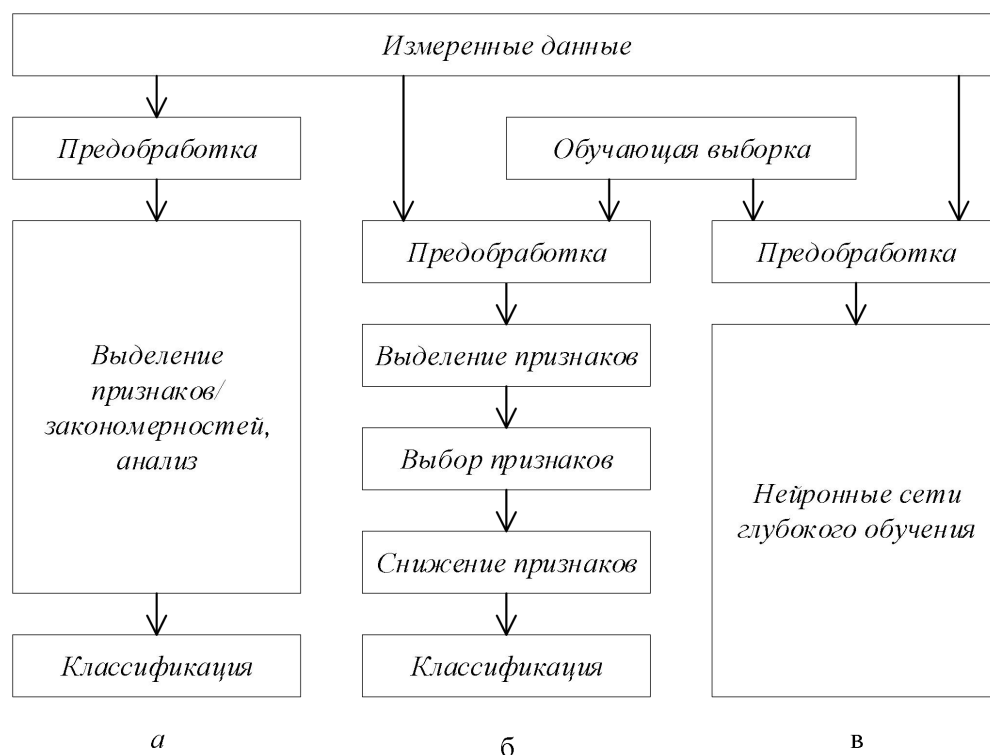


Рисунок 2. Этапы работы алгоритмов распознавания аритмий: *а* – классические алгоритмы без обучения; *б* – алгоритмы машинного обучения; *в* – алгоритмы глубокого машинного обучения

Общий подход к предварительной обработке данных. ЭКГ является относительно «сильным» сигналом с легко различимыми волнами. Большинство типов помех, которые воздействуют на сигнал ЭКГ, могут быть устранены с помощью полосовых фильтров [5]. Этот шаг предварительной обработки делает сигнал ЭКГ пригодным для последующих процессов. Наиболее распространенные методы, используемые для уменьшения помех, следующие:

- низкочастотные и высокочастотные фильтры Баттерворта;
- вейвлеты Добеши;
- ортогональные вейвлет-фильтры.

Также анализ сигнала для мониторингового наблюдения или диагностики требует идентификации эпох и исследования соответствующих событий. После того, как было идентифицировано какое-либо событие, осуществляется сегментация и анализ соответствующей волны сигнала с использованием таких характеристик, как амплитуда, форма волны (морфология), длительность, интервалы между событиями, распределение энергии, частотное содержание и т. д. Форма и амплитуда QRS-комплекса, как наиболее значимые параметры ЭКГ, являются предметом исследования при любой диагностике и анализе биений сердца. Пан и Томпкинс (*Pan and Tompkins*) [16, 17] предложили алгоритм обнаружения QRS-комплексов, рассчитанный на работу в режиме реального времени и основанный на анализе наклона, амплитуды и ширины QRS-комплекса.

Общий подход к распознаванию и классификации образов. Распознавание или классификацию образов можно определить, как отнесение входных данных к какому-либо из идентифицируемых классов посредством дифференциации существенных признаков или атрибутов данных от фонового сигнала и от деталей, не относящихся к изучаемому явлению. После того, как из сигнала были получены количественные признаки, он может

быть представлен вектором признаков $x = (x_1, x_2, \dots, x_n)^T$, известным также как вектор параметров или вектор образца. Когда величины x_i являются вещественными числами, вектор x представляет собой точку в n -мерном евклидовом пространстве. Ожидается, что векторы сходных объектов формируют кластеры.

Для эффективной классификации образов необходимы такие признаки, которые могут дать непересекающиеся наборы или кластеры для векторов параметров. Это обстоятельство подчеркивает важность правильного выбора процедур предварительной обработки и выделения признаков. Признаки или свойства, являющиеся общими для всех образцов, принадлежащих к определенному классу, известны как внутрикласовые признаки. Дискриминирующие признаки, которые характеризуют различия между классами образов, называются межклассовыми признаками.

Проблема классификации образов заключается в определении разделяющих границ или разработке решающих процедур, для отнесения данных к различным классам образов, основанных на векторах признаков [5].

Выделение признаков. Выделение признаков представляет собой интерактивный процесс, включающий ряд автоматических процедур преобразования данных. При большом количестве измерений-признаков, описывающих характеристики входного сигнала, для сокращения размерности задачи применяется корреляционный и факторный анализ данных. По методу извлечения и методу анализа, признаки можно разделить на следующие категории:

Временные и морфологические признаки. Эти признаки описаны во временной области и представляют из себя амплитуду, наклон и частоту сердечных сокращений (ЧСС). Наиболее частыми характеристиками являются *интервал пересечения порога (TCI)* [18], количество отсчетов пересечения порога (TCSC) [19], *среднее абсолютное значение (MAV)* [20], *стандартная экспонента (STE)* и *модифицированная экспонента (MEA)* [21].

Спектральные признаки. Эти признаки определены в частотной области. Они могут учитывать плотность спектра, нормированные спектральные моменты, которые характеризуют распределение точек спектра, и соответствующая информация о мощности спектра в разных диапазонах частот. Эти алгоритмы экстракции признаков включают *утечку фильтра VF (VFleak)* [22], спектральный алгоритм [23] и алгоритм *медианной частоты (MF)* [24]. *VFleak* – это функция отклика узкополосного полосового фильтра, которая определяет среднюю частотную область сегмента ЭКГ и его выход – утечку фильтра *VF*. Спектральный алгоритм вычисляет информацию о мощности и энергии по разным частотным диапазонам с использованием Фурье-анализа. *Медианная частота* – это центральная частота на распределении частот, найденных в заданном сегменте ЭКГ.

Выбор признаков. Этот шаг позволяет удалить избыточные функции, снизить вычислительные затраты и улучшить общую производительность системы. На этом этапе используются три основные категории методов выбора признаков:

1 Методы оболочки: считаются лучшим подходом для выбора признаков с точки зрения точности, но обладают значительной вычислительной сложностью [25]. В этих методах используется перекрестная проверка путем многократного обучения модели с использованием различных признаков и сравнение результатов. Распространены следующие методы:

- рекурсивное исключение признаков;
- прямой выбор признаков;
- генетические алгоритмы;

2 Методы фильтры: этот подход использует статистические измерения для выбора наилучшего набора характеристик перед процессом обучения. Здесь функция оценивается по методу прокси, а не по точности перекрестной проверки. Используются следующие методы:

- корреляция;
- хи-квадрат;
- дисперсионный анализ;
- ReliefF*;

3 Встроенные методы: данный подход включает методы, которые не подпадают под вышеупомянутые подходы (оболочка или фильтр). *L1*-регуляризация является примером такого метода.

Снижение признаков. На данном этапе выполняется уменьшение числа признаков набора данных. Многие из методов снижения признаков преобразуют выбранные признаки в пространства с пониженной размерностью. Анализ главных компонент, линейный дискриминантный анализ и локально-чувствительный дискриминантный анализ являются примерами методов снижения признаков.

Классификация. Классификация – это основной блок распознавания для решаемой задачи, хотя отдельные процедуры распознавания могут присутствовать и на этапе выделения признаков, и на этапе постклассификационной обработки. Но именно результат, полученный на этапе классификации, является основой для интерпретации сигнала ЭКГ как шокового или нешокового. Процесс классификации может осуществляться в несколько этапов или в несколько итераций, в зависимости от выбранной схемы распознавания. В ряде случаев результаты, полученные на данном этапе, требуют пересмотра всей схемы обработки в целом. Наиболее распространены методы классификации *метод опорных векторов (SVM)*, *наивный баесовский классификатор*, *метод анализа k-ближних соседей (k-NN)*, *деревья решений (DT)*.

Методы глубокого машинного обучения. В последние годы применение глубоких нейронных сетей, вызвали большой интерес в компьютерной диагностике практически всех заболеваний [26; 27]. В зависимости от архитектуры сети глубокого обучения подразделяются на различные классы, такие как полностью связанные сети, сети убеждений и сверточные сети. Одним из основных преимуществ такого подхода является то, что они выполняют извлечение признаков автоматически. Наиболее распространенные архитектуры, используемые в сигнале ЭКГ, – это *сверточная нейронная сеть (CNN)* и *рекуррентная нейронная сеть (RNN)*.

Создание модели. Для создания модели используется база данных ЭКГ *MIT-BIH Arrhythmia Database* [8]. Источником данных ЭКГ, включенных в эту базу аритмий, является набор из более чем 4000 долговременных записей Холтера, которые были получены лабораторией аритмии в больнице Бет Исраэль в период с 1975 по 1979 год. Примерно 60% этих записей были получены от пациентов, находящихся в стационаре. База данных содержит 23 записи (группа 1), выбранные случайным образом из этого набора, и 25 записей (группа 2), выбранные из того же набора для включения разновидности редких, но клинически важных явлений, которые не могут быть хорошо представлены небольшой случайной выборкой записей Холтера. Каждая из 48 записей длится чуть более 30 минут.

Первая группа записей призвана служить репрезентативной выборкой различных форм сигналов и артефактов, с которыми детектор аритмии может столкнуться при обычном клиническом использовании. Ленты с записями отбирались случайным образом для выделения из них получасовых отрезков. Выбранные таким образом сегменты исключались только в том случае, если ни один из двух сигналов ЭКГ не имел достаточного качества для анализа людьми-экспертами.

Записи во второй группе были выбраны для включения сложных желудочковых, узловых и наджелудочковых аритмий и нарушений проводимости. Некоторые из этих записей были выбраны потому, что особенности ритма, вариаций морфологии *QRS* или

качества сигнала могут представлять значительные трудности для детекторов аритмии; эти записи приобрели значительную известность среди пользователей баз данных.

Испытуемыми были 25 мужчин в возрасте от 32 до 89 лет и 22 женщины в возрасте от 23 до 89 лет.

Для работы с записями базы данных используется фреймворк *wfdb-python* [28], который позволяет считывать записи (рисунок 3) и работать с ними как с массивами данных.

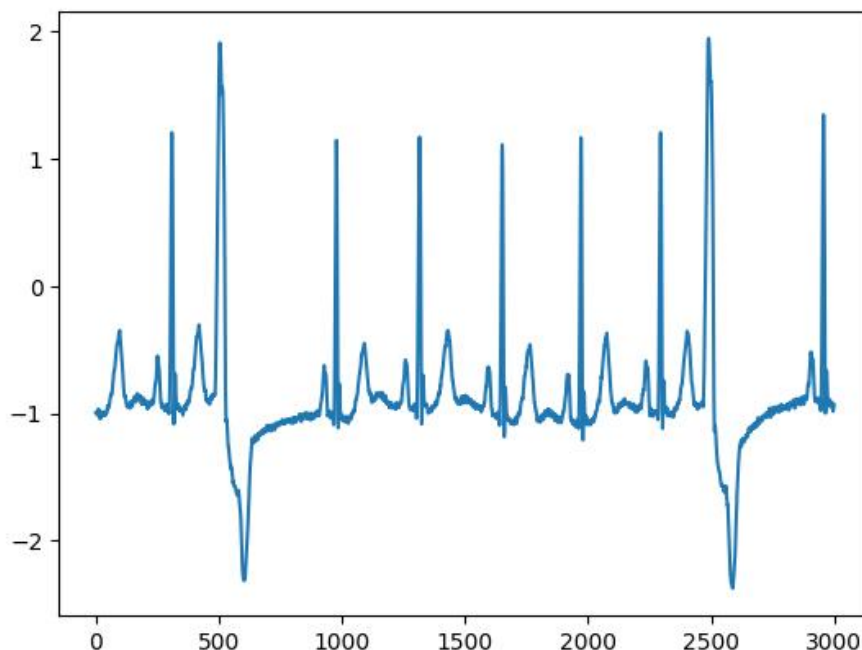


Рисунок 3. Вид сигнала ЭКГ одной из записей базы данных

Первым шагом в подготовке данных является извлечение *RR-интервалов* из записей базы данных для определения *вариабельности сердечного ритма (BCP)*, где *R* – точка, соответствующая пику комплекса *QRS* волны ЭКГ и *R-R* – интервал между последовательными *R* (рисунок 4).

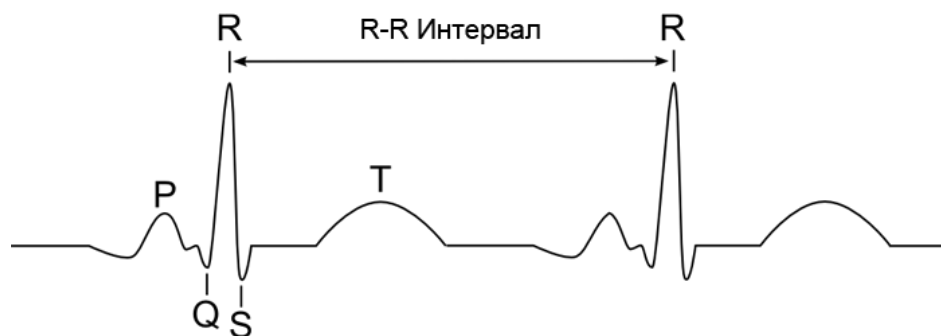


Рисунок 4. Представление *QRS*-комплекса и *RR*-интервала на ЭКГ

После обработки всех записей получается массив данных, который содержит все *RR* интервалы из базы данных. Полученные интервалы разбиваются на сегменты по 10 записей в каждом и определяется количество аномальных ударов. Затем эти данные разделяются на две подгруппы, одна из которых будет использоваться для обучения, а вторая – для проверки работоспособности обученной модели. Метки, присутствующие в

записях, позволяют однозначно определить класс обнаруженного события. Для обучения модели используется фреймворк *TensorFlow* [29]. Результаты обучения модели [30] представлены на рисунке 5.

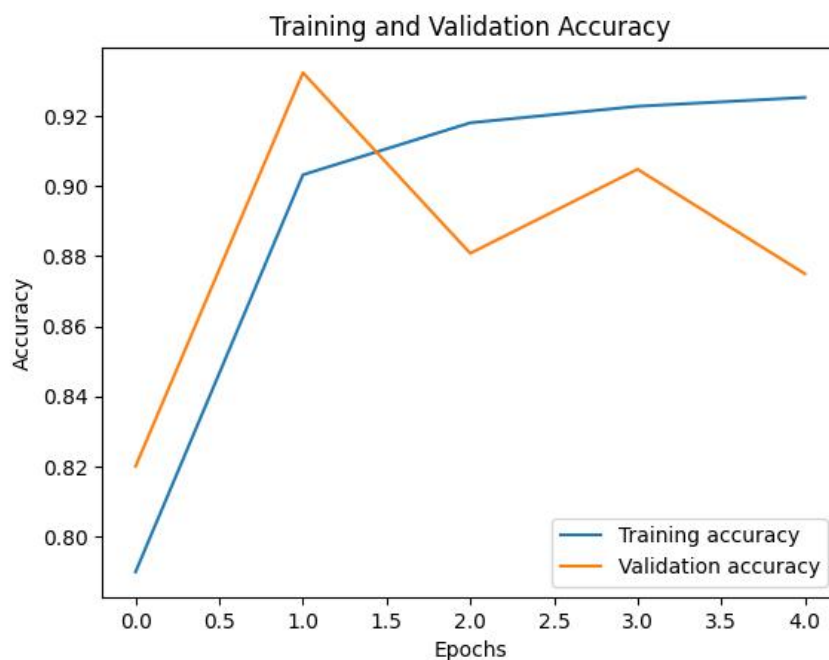


Рисунок 5. Точность обучения и проверки модели

Заключение. Рассмотрены методы автоматического анализа ЭКГ с целью получения предупреждения, что в сигнале ЭКГ есть аномалии. На примере тестового набора из *MIT-BIH Arrhythmia Database* оценена точность работы модели [30]. Достигнутая точность составила 92%. Полученную модель планируется использовать во встроенной системе совместно с фреймворком *TinyML* [31] для анализа сигнала ЭКГ в режиме реального времени на наличие аритмии.

Список литературы

- [1] Cardiovascular diseases [Электронный ресурс] / World Health organization. – Режим доступа: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1. – Дата доступа: 11.02.2024.
- [2] Аритмия сердца [Электронный ресурс] / Wikipedia. – Режим доступа: https://ru.wikipedia.org/wiki/Аритмия_сердца. – Дата доступа: 11.02.2024.
- [3] Суточное мониторирование ЭКГ [Электронный ресурс] / Wikipedia. – Режим доступа: https://ru.wikipedia.org/wiki/Суточное_мониторирование_ЭКГ. – Дата доступа: 11.02.2024.
- [4] Kennedy H. The Evolution of Ambulatory ECG Monitoring / H. Kennedy // *Progress in Cardiovascular Diseases*. – 2013. – Vol. 56. – No. 2. – P. 127-132.
- [5] Рангайян Р.М. Анализ биомедицинских сигналов. Практический подход / Пер. с англ. под ред. А.П. Немирко. – М.: ФИЗМАЛИТ, 2010. – 440 с.
- [6] Flores N. Readily available ECG databases / N. Flores, R.L. Avitia, M.A. Reyna, C. Garcí // *Journal of Electrocardiology*. – 2018. – Volume 51. – P.1095-1097. – DOI: 10.1016/j.jelectrocard.2018.09.012.
- [7] PTB Diagnostic ECG Database [Электронный ресурс] / PhysioNet. – Режим доступа: <https://www.physionet.org/content/ptbdb/1.0.0/>. – Дата доступа: 11.02.2024.
- [8] MIT-BIH Arrhythmia Database [Электронный ресурс] / PhysioNet. – Режим доступа: <https://www.physionet.org/content/mitdb/1.0.0/>. – Дата доступа: 11.02.2024.
- [9] St Petersburg INCART 12-lead Arrhythmia Database [Электронный ресурс] / PhysioNet. – Режим доступа: <https://physionet.org/content/incartdb/1.0.0/>. – Дата доступа: 11.02.2024.
- [10] MIT-BIH Supraventricular Arrhythmia Database [Электронный ресурс] / PhysioNet. – Режим доступа: <https://physionet.org/content/svdb/1.0.0/>. – Дата доступа: 11.02.2024.

- [11] АНА Database [Электронный ресурс] / ECRI. Режим доступа: <https://www.ecri.org/>. – Дата доступа: 11.02.2024.
- [12] UofT ECG Database [Электронный ресурс] / BioSec.lab. – Режим доступа: <https://www.comm.utoronto.ca/~biometrics/databases.html>. – Дата доступа: 11.02.2024.
- [13] Telemetric and Holter ECG Warehouse [Электронный ресурс] / THEW. – Режим доступа: <http://thew-project.org/>. – Дата доступа: 11.02.2024.
- [14] PHYSIOBC [Электронный ресурс] / physiobc.org. – Режим доступа: <http://www.physiobc.org/>. – Дата доступа: 11.02.2024.
- [15] Липчак Д.А. Обзор методов автоматической диагностики сердечной аритмии для принятия решений о необходимости проведения дефибрилляции / Д.А. Липчак, А.А. Чупов // *Ural Radio Engineering Journal*. – 2021. – №5(4). – С.380-409. – DOI: 10.15826/urej.2021.5.4.004.
- [16] Pan J. A Real-Time QRS Detection Algorithm / J. Pan, W.J. Tompkins // *IEEE Transactions on Biomedical Engineering*. – 1985. – P. 230-236.
- [17] Clifford G.D. Signal processing methods for heart rate variability analysis. PhD Thesis. Michaelmas Term. 2002. 244 p.
- [18] Thakor N.V. Ventricular Tachycardia and fibrillation detection by a sequential hypothesis testing algorithm / N.V. Thakor, Y.S. Zhu, K.Y. Pan // *IEEE Transactions on Biomedical Engineering*. – 1990. – №37(9). – P.837-843. – DOI:10.1109/10.58594.
- [19] Arafat M.A. A simple time domain algorithm for the detection of ventricular fibrillation in electrocardiogram / M.A. Arafat, A.W. Chowdhury, M.K. Hasan // *Signal, Image and Video Processing*. – 2011. – №5(1). – P.1-10. – DOI: 10.1007/s11760-009-0136-1.
- [20] Anas E. Sequential algorithm for life threatening cardiac pathologies detection based on mean signal strength and EMD functions / E. Anas, S.Y. Lee, M.K. Hasan // *BioMedical Engineering OnLine*. – 2010. – №9(1). – DOI:10.1186/1475-925X-9-43.
- [21] Amann A. Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators / A. Amann, R. Tratnig, K. Unterkofler // *BioMedical Engineering OnLine*. – 2005. – №4:60. – DOI: 10.1186/1475-925X-4-60.
- [22] Granitto P. M. Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products / P.M. Granitto, [etc.] // *Chemometrics and Intelligent Laboratory Systems*. – 2006. – №83(2). – P.83-90. – DOI: 10.1016/j.chemolab.2006.01.007.
- [23] Barro S. Algorithmic sequential decision-making in the frequency domain for life threatening ventricular arrhythmias and imitative artefacts: a diagnostic system / S. Barro, [etc.] // *Journal of Biomedical Engineering*. – 1989. – №11(4). – P.320-328. – DOI: 10.1016/0141-5425(89)90067-8.
- [24] Pławiak P. Novel genetic ensembles of classifiers applied to myocardium dysfunction recognition based on ECG signals / P. Pławiak // *Swarm and Evolutionary Computation*. – 2018. – №39. – P.192-208. – DOI: 10.1016/j.swevo.2017.10.002.
- [25] Blum A.L. Selection of relevant features and examples in machine learning / A.L. Blum, P. Langley // *Artificial Intelligence*. – 1997. – №97(1-2). – P.245-271. – DOI:10.1016/S0004-3702(97)00063-5.
- [26] Miotto R. Deep learning for healthcare: Review, opportunities and challenges / R. Miotto, [etc.] // *Briefings in Bioinformatics*. – 2018. – №19(6). – P.1236-1246. – DOI: 10.1093/bib/bbx044.
- [27] Esteva A. A guide to deep learning in healthcare / A. Esteva, [etc.] // *Nature Medicine*. – 2019. – №25(1). – P.24-29. – DOI: 10.1038/s41591-018-0316-z.
- [28] WFDB-python [Электронный ресурс] / Native Python WFDB package. – Режим доступа: <https://github.com/MIT-LCP/wfdb-python>. – Дата доступа: 11.02.2024.
- [29] TensorFlow [Электронный ресурс] / An end-to-end machine learning platform. – Режим доступа: <https://www.tensorflow.org/>. – Дата доступа: 11.02.2024.
- [30] TinyRhythmAnalyzer [Электронный ресурс] // TinyML-powered ECG arrhythmia detection on Arduino Nano 33 BLE Sense. – Режим доступа: <https://github.com/tonyfu97/TinyRhythmAnalyzer/tree/main/model>. – Дата доступа: 11.02.2024.
- [31] TinyML [Электронный ресурс] // tinyML Foundation. – Режим доступа: <https://www.tinyml.org/>. – Дата доступа: 11.02.2024.

Авторский вклад

Реджепов Владимир Александрович – описание методов автоматической диагностики аритмии, тестирование модели нейронной сети, формирование структуры статьи.

Перцев Дмитрий Юрьевич – руководство исследованием, анализ полученных результатов.

AUTOMATIC DIAGNOSTICS OF CARDIAC ARRHYTHMIA IN EMBEDDED SYSTEMS

V. Redzhepov

Post-graduate student,

*Department of Electronic Computing Machines
of BSUIR*

D. Pertsau

Associate Professor, PhD,

*Department of Electronic Computing Machines
of BSUIR*

Abstract. The technological development of wearable devices has led to the emergence of simple and compact embedded ECG monitoring systems. At the same time, result analysis is one of the key performance indicators for commercial monitoring systems. This is because classification, as one element of the analysis, makes it possible to identify an abnormal ECG signal and therefore propose appropriate treatment. Since this requires computing power, a client-server approach is usually considered. This means that the embedded system is only a data recording and storage system in a traditional ECG monitoring system. However, to enable real-time diagnosis, the need for ECG classification in embedded systems has been introduced. This paper reviews existing methods for automatic diagnosis of cardiac arrhythmias and proposes a solution for identifying cardiac arrhythmias in embedded systems.

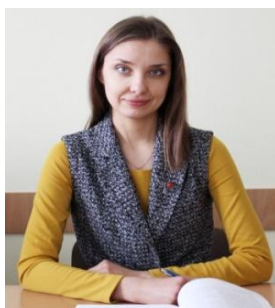
Keywords: ECG, arrhythmia, digital signal processing, real time, machine learning, classification, TinyML, embedded systems.

УДК 621.3.049.77–048.24:537.2

АНАЛИЗ ДИНАМИЧЕСКИХ ХАРАКТЕРИСТИК ДВИЖЕНИЙ СПОРТСМЕНОВ С ИСПОЛЬЗОВАНИЕМ ДИНАМОМЕТРИЧЕСКОЙ ПЛАТФОРМЫ



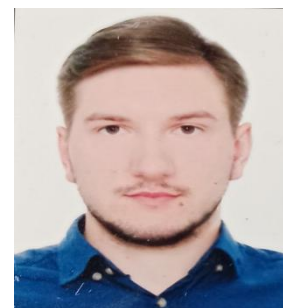
М.В. Давыдов
Кандидат технических наук, доцент, первый проректор БГУИР
davydov-mv@bsuir.by



Н.С. Давыдова
Кандидат технических наук, доцент кафедры инфокоммуникационных технологий, доцент
davydovans@bsuir.by



Д.Ю. Тербиленко
Магистрант факультета компьютерного проектирования БГУИР, специальности «Электронные системы и технологии»
daniil.terbilenko@gmail.com



Н.М. Елец
Магистрант факультета компьютерного проектирования БГУИР, специальности «Электронные системы и технологии»
eletz.nikita@yandex.by

М.В. Давыдов

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием алгоритмов цифровой обработкой биомедицинских сигналов и изображений, моделирование биологических объектов, разработка программных продуктов медицинского и спортивного назначения.

Н.С. Давыдова

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием сетевых технологий, микропроцессорных устройств, цифровой обработкой сигналов, спортивной и биомедицинской инженерией.

Д.Ю. Тербиленко

Окончил Белорусский государственный университет информатики и радиоэлектроники.

Н.М. Елец

Окончил Белорусский государственный университет информатики и радиоэлектроники.

Аннотация. Проведен анализ технических компонентов вертикального прыжка. Предложены алгоритмы автоматизированной оценки динамических характеристик движений спортсменов на основании анализа данных динамометрической платформы. Представлены результаты разработки программного обеспечения для анализа и визуализации данных динамометрической платформы.

Ключевые слова: динамометрическая платформа, динамические характеристики движений, цифровая обработка сигналов.

Введение. В области спортивных исследований существует проблема автоматизации обработки данных. В частности, актуальным является разработка алгоритмов автоматизированного анализа динамических характеристик движений человека.

Динамометрические платформы широко используются в различных областях, включая автомобильные краш-тесты, клинический анализ походки и спортивной техники. Платформа представляет собой прямоугольную металлическую пластину, на которой расположены пьезоэлектрические или тензометрические датчики. Датчики регистрируют физические силы, действующие на пластину, и генерируют электрический выходной сигнал, пропорциональный этим силам [1, 2].

Динамометрическая платформа может использоваться для анализа динамических характеристик движений спортсменов. К ним относятся ускорение, скорость, перемещение общего центра масс и др. [3]. Для расчёта этих параметров используется специальное программное обеспечение для регистрации и визуализации вертикальной составляющей силы реакции опоры и смещения общего центра масс спортсмена. Анализ данных дает эффективную иллюстрацию соотношений между силами, действующими на тело спортсмена, и результирующими ускорением, скоростью и перемещением тела.

Выбор исследуемых движений спортсменов. В представленной работе для анализа динамических характеристик движений спортсменов выбран вертикальный прыжок. Вертикальный прыжок – это атлетическое действие, целью которого является максимальное возвышение общего центра масс тела (ОЦМ) в вертикальном направлении относительно положения покоя атлета. Эффективность прыжка определяется силой, приложенной атлетом во время отталкивания [4].

Во время вертикального прыжка равнодействующая сила, действующая на общий центр масс спортсмена, рассчитывается по формуле (1):

$$F = F_{CPO} - mg, \quad (1)$$

где F_{CPO} – сила реакции опоры, действующая на спортсмена, Н;

m – масса спортсмена, кг;

g – ускорение свободного падения, m/c^2 .

Именно ее величина и вектор определяет динамические характеристики атлета во время вертикального прыжка (рисунок 1). Она может быть как положительной (сила реакции опоры больше силы тяжести), так и отрицательной (сила реакции опоры меньше силы тяжести). Направление силы влияет и на направление ускорения. Если сила/ускорение с положительным знаком – атлет ускоряется вверх (встает), если с отрицательным знаком – атлет ускоряется вниз (падает).

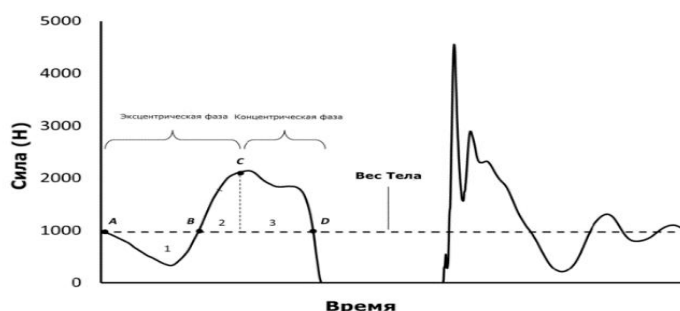


Рисунок 1. Динамометрическая кривая вертикального прыжка

В работе исследовались различные виды вертикального прыжка: *Countermovement jump*, *Square jump* и *Abalakov jump* [5].

Расчёт основных динамических характеристик движений. На основании результирующей силы спортсмена (динамометрической кривой) проведем расчет основных динамических характеристик движения:

1 Ускорение движения (формула 2):

$$a_i = (F_{zi} - P)/m, \quad (2)$$

где F_{zi} – вертикальная составляющая результирующей силы в момент времени i ;
 P – сила тяжести.

Пример кривой ускорения для прыжка *Countermovement jump* представлен на рисунке 2.

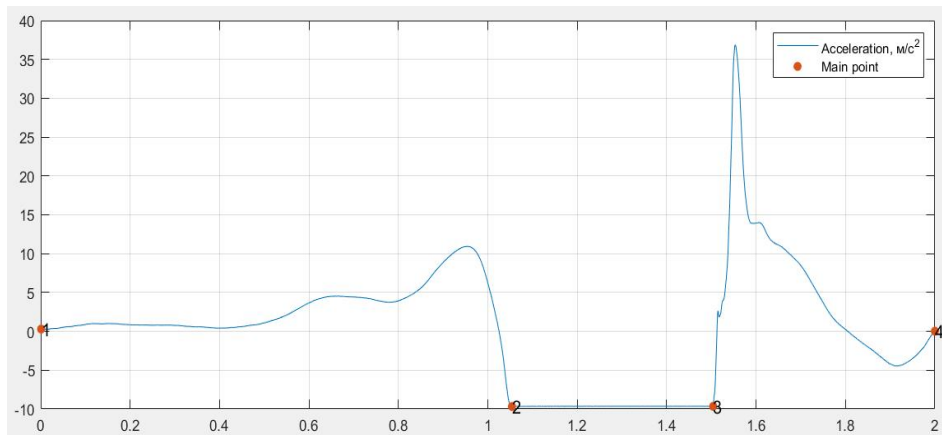


Рисунок 2. Кривая ускорения для вертикального прыжка

2 Скорость движения (формула 3):

$$V_i = V_{i-1} + a_i * dt, \quad (3)$$

где V_i – скорость движения в момент времени i ;
 dt – время дискретизации.

Пример кривой скорости для прыжка *Square jump* представлен на рисунке 3.

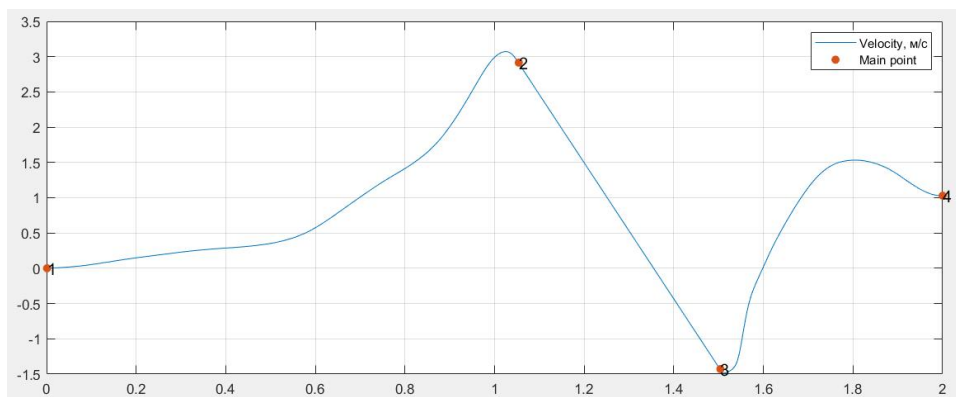


Рисунок 3. Кривая скорости для вертикального прыжка

3 Перемещение ОЦМ (формула 4):

$$D_i = D_{i-1} + V_{i-1} * dt + a_i * dt^2 / 2, \quad (4)$$

где D_i – перемещение в момент времени i .

Пример кривой перемещения ОЦМ для прыжка *Abalakov jump* представлен на рисунке 4.

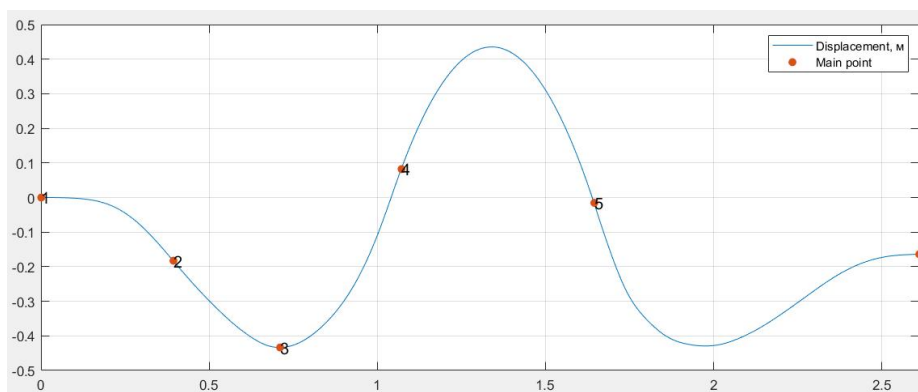


Рисунок 4. Кривая перемещения общего центра масс для вертикального прыжка

4 Импульс движения (формула 5):

$$p_i = p_{i-1} + (F_{z_i} - P) * dt, \quad (5)$$

где p_i – импульс движения в момент времени i .

5 Мощность движения (формула 6):

$$N_i = F_{z_i} * V_i, \quad (6)$$

где N_i – мощность движения в момент времени i .

6 Фазовый портрет движения (зависимость силы движения от перемещению ОЦМ) для прыжка *Countermovement jump* представлен на рисунке 5.

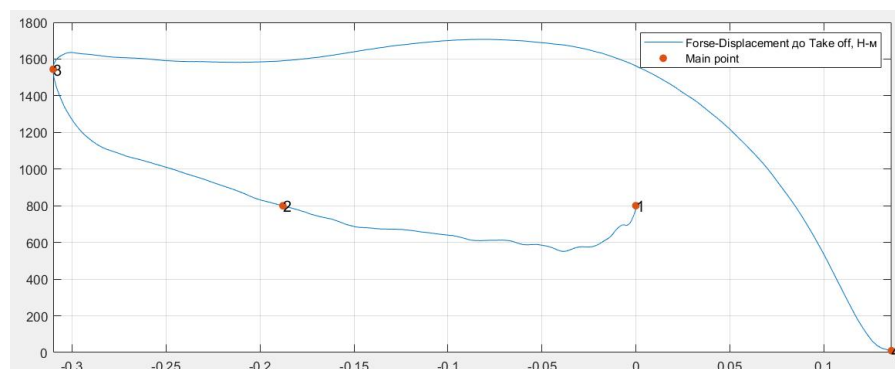


Рисунок 5. Фазовый портрет движения

Разработка программного средства для анализа динамических характеристик движения спортсменов. Основные алгоритмы обработки данных динамометрической платформы выполнены в пакете прикладных программ для решения задач технических вычислений *MATLAB*.

Программа для анализа и визуализации данных динамометрической платформы написана на языке *MATLAB* с использованием среды разработки *MATLAB App Designer* [6].

Для обеспечения удобства взаимодействия пользователя с программой, ее возможности были разделены на несколько взаимосвязанных окон:

1 Основное окно программы (рисунок 6).

В основном окне программы реализованы следующие функции:

– Отображение записей спортсменов, хранящихся в базе данных.

– Возможность добавления, редактирования и удаления записей спортсменов.

- Фильтрация отображения записей спортсменов по спортивной квалификации.
- Сортировка отображения записей по убыванию или возрастанию.
- Выбор типа прыжка загружаемого файла тензометрической установки.
- Отображение результатов анализа ранее обработанных записей прыжков для каждого спортсмена хранящихся в базе данных с возможностью удаления.

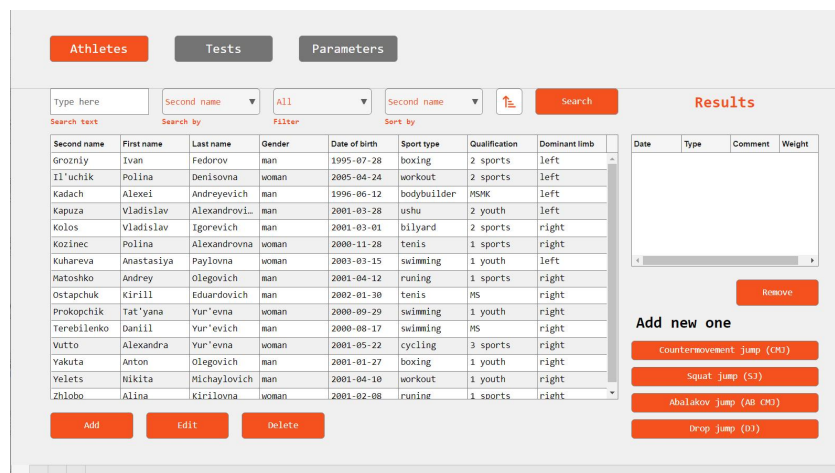


Рисунок 6. Основное окно программы для анализа и визуализации данных динамометрической платформы

2 Окно предварительной обработки данных динамометрической платформы.

После выбора файла для анализа, программа автоматически открывает окно, в котором отображается график вертикальной силы реакции опоры. С использованием разработанных и проверенных алгоритмов в *MATLAB*, программа также автоматически определяет границы каждого прыжка, так как в одном файле записывается несколько попыток движения. Результатом анализа данных также является определение массы спортсмена и построение изолинии, соответствующей его весу (рисунок 7).

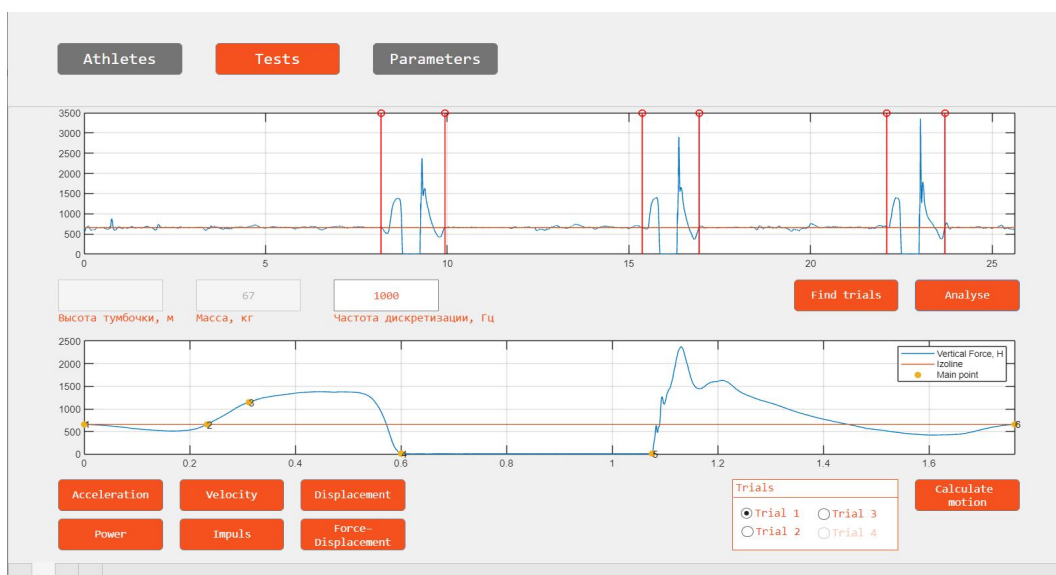


Рисунок 7. Окно предварительной обработки данных динамометрической платформы

3 Окно расчёта динамических характеристик движения спортсмена.

После нажатия на кнопку расчёта параметров движения, откроется окно, в котором можно выбрать интересующие динамические характеристики движения для последующего расчёта, а также сохранить рассчитанные характеристики в EXCEL файл (рисунок 8).

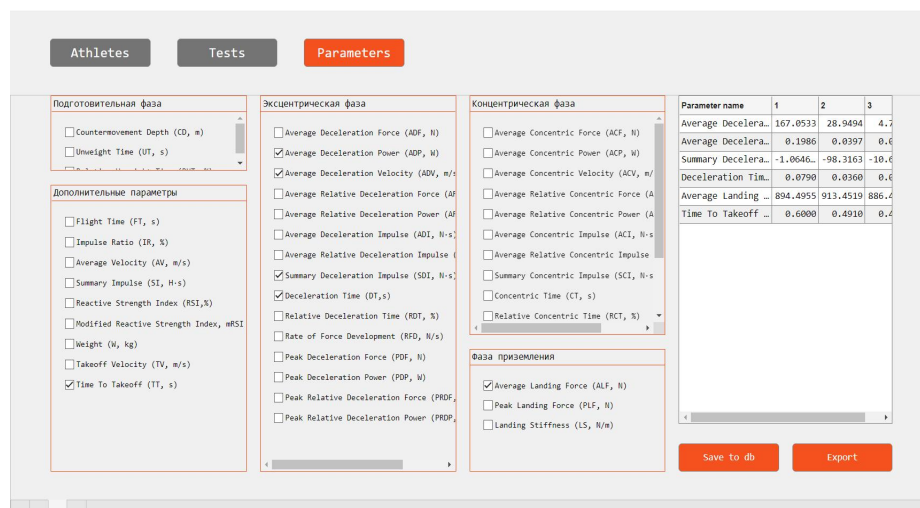


Рисунок 8. Окно расчёта динамических характеристик движения спортсмена

Динамические параметры движения рассчитываются на основании ранее построенных сигналов и выделенных фаз движения. Расчет производится для каждого выделенного цикла движения, а затем осуществляется статистический анализ динамических параметров для всех выделенных циклов движения (математическое ожидание, СКО, максимальное и минимальное значение). Всего в программе рассчитывается более 40 показателей.

Закключение. Разработанные в ходе работы алгоритмы и программное средство для автоматизированного анализа динамических характеристик движений спортсменов на основе динамометрической платформы могут быть использованы для оценки спортивного мастерства, контроля результатов тренировочного процесса, прогнозирования двигательной одаренности у детей и эффективном обучении двигательным навыкам.

Дальнейшее развитие программного обеспечения планируется в разрезе детализации каждого отдельного прыжка, кроссплатформенности и размещении базы данных на удаленном сервере, что позволит получать доступ к записям спортсменов с разных устройств.

Список литературы

- [1] R. Cross, «Standing, walking, running, and jumping on a force plate» Am. J. Phys. 67(4), 304–309 (1998).
- [2] J. A. Major, W. A. Sands, J. R. McNeal, D. D. Paine, and R. Kipp, «Design, construction, and validation of a portable one-dimensional forceplatform» J. Strength Conditioning Res. 12(1), 37–41 (1998).
- [3] Загrevский, В. И. Биомеханика физических упражнений : учебное пособие / В. И. Загrevский, О. И. Загrevский – Томск : Издательский дом Томского государственного университета, 2018. – 262 с.
- [4] Вертикальный прыжок: силы (часть 1) [Электронный ресурс]. – Режим доступа: <https://zen.yandex.ru/media/abulahov/vertikalnyi-pryjok-sily-chast-1-5ba1028c14b74200aefe005e> – Дата доступа: 11.11.2023.
- [5] P. V. Komi, “Stretch-shortening cycle,” in Strength and Power in Sport, edited by P. V. Komi (Blackwell Science, Oxford, 1992), pp. 169–179.
- [6] Develop Apps Using App Designer [Электронный ресурс]. – Режим доступа : <https://www.mathworks.com/help/matlab/app-designer.html> – Дата доступа: 14.11.2023.

Авторский вклад

Давыдов Максим Викторович – анализ и постановка задачи, общий анализ алгоритмов цифровой обработки сигналов

Давыдова Надежда Сергеевна – разработка алгоритмов автоматизированного анализа данных динамометрической платформы.

Теребиленко Даниил Юрьевич – разработка графического интерфейса пользователя, создание структуры базы данных.

Елец Никита Михайлович – разработка десктопной программы анализа с графическим интерфейсом.

ANALYSIS OF THE DYNAMIC CHARACTERISTICS OF ATHLETES' MOVEMENTS USING A DYNAMOMETRIC PLATFORM

M.V. Davydov

*PhD of Technical Sciences,
Associate Professor
First Vice-Rector*

Davydova N.S.

*PhD of Technical Sciences,
Associate Professor of the
Department of
Infocommunication
Technologies, Associate
Professor*

D.Y. Terebilenko

*Master's student at the
Faculty of Computer
Science
design BSUIR, specialty
«Electronic systems and
technologies»*

N.M. Yelets

*Master's student at the
Faculty of Computer
Science
design BSUIR, specialty
«Electronic systems and
technologies»*

Abstract. An analysis of the technical components of the vertical jump was carried out. Algorithms for automated assessment of the dynamic characteristics of athletes' movements are proposed based on the analysis of data from a dynamometer platform. The results of software development for analysis and visualization of dynamometer platform data are presented.

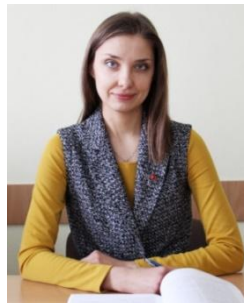
Key words: dynamometer platform, dynamic characteristics of movements, digital signal processing.

УДК 004.021:004.622:004.043

КИНЕМАТИЧЕСКАЯ МОДЕЛЬ ДВИЖЕНИЙ ЧЕЛОВЕКА НА БАЗЕ МАТРИЦ РАССТОЯНИЙ



Д.И. Гусейнов
Аспирант кафедры
«Спортивная инженерия БНТУ»
guseynov.daniil@yandex.by



Н.С. Давыдова
Доцент кафедры
инфокоммуникационных
технологий, кандидат
технических наук, доцент
davydova-ns@bsuir.by

Д.И. Гусейнов

Окончил Белорусский национальный технический университет. Область научных интересов связана с разработкой методов и алгоритмов обработки, автоматизированного анализа и оценки данных.

Н.С. Давыдова

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием сетевых технологий, микропроцессорных устройств, исследование особенностей цифровой обработки сигналов, организацией учебного и научно-исследовательского процессов в техническом университете.

Аннотация. В статье рассмотрен подход анализа и контроля техники спортивных движений с использованием данных системы захвата движений и нормированных матриц расстояний. Выполнен анализ кинематической структуры тела спортсмена, а также некоторые аспекты формирования содержания нормированных матриц расстояний. Приведен матрицы расстояний для ключевых поз спортсмена в технике бега на 400 м.

Ключевые слова: системы захвата движений, матрицы, нормализация, кинематика, модель.

Введение. Системы захвата движений, в основе которых лежат различные вычислительные технологии, используются во многих сферах жизнедеятельности человека с целью поддержки принятия решения, предоставления необходимого и содержательного объема достоверных данных, обеспечения безопасности, автоматизация промышленных и гражданских процессов [1]. Наиболее насыщенными технологиями захвата движений областями является строительство, робототехника и автомобилестроение [2]. В силу своих функциональных и вычислительных возможности системы захвата движений все чаще используются в других областях. Одними из наиболее перспективных областей для использования подобных систем являются область спортивной тренировки и спортивных исследований.

Наиболее актуальными и труднорешаемыми задачами, формируемыми в рамках современной спортивной тренировки вне зависимости от вида спорта, являются задачи, связанные с контролем динамики подготовленности, повышение его двигательной производительности, совершенствование техники выполнения двигательных действий, снижение риска травматизма, а также, что в значительной степени актуально в последнее

время, формирование и предоставление спортсмену объективной и содержательной обратной связи [3]. К задачам, непосредственно решаемым в области спортивных исследований, относятся изучение биомеханических механизмов организации двигательного действия, комплексный анализ и синтез целостного упражнения или отдельных его частей [4]. Все обозначенные задачи в необходимой и достаточной степени решаются с использованием систем захвата движений.

Несмотря на тот факт, что системы захвата движений могут применяться в самых различных областях, а предоставляемые ими данные могут быть использованы при решении задач различной сложности и комплексности, главной проблемой, ограничивающей использование таких систем, является формат и структура представления данных для дальнейшей их предметной обработки и анализа. В частности, в области спортивной биомеханики часто рассматривается тело спортсмена как структурная совокупность биокинематических звеньев и сочленений. Часто такая структурная совокупность представляет собой большой содержательный набор данных. В данном случае необходимо использовать такой формат и структуру данных, которая позволила бы сохранить полноту биомеханических признаков, характерных для того или иного двигательного действия. Наиболее подходящей структурой данных, позволяющей сохранить такие признаки с учетом широты регистрируемой и обрабатываемой информации, а также однозначно количественно описать двигательное действие, является матрица расстояний для каждой позы в ключевые моменты времени, а в качестве формата данных, наполняющих матрицу, – нормализованные показатели пространственного расстояния.

Кинематическая модель тела спортсмена. Базовая модель тела человека при решении прикладных задач спортивной биомеханики структурно состоит из трех звеньев – ноги, туловище, руки. Однако подобная модель является плоскостной, не обладает достаточной детализацией для анализа сложных многоплоскостных двигательных действий, сочетающих в себе поступательное и вращательное движение, а также характеризуется наличием определенных критических анатомических ограничений [5, 6]. В современной спортивной биомеханике в рамках решения задач, связанных с количественным анализом и оценкой двигательных действий спортсмена, вне зависимости от вида спорта и пространственной сложности упражнения, используется формализованная модель тела человека, которую можно представить в виде упорядоченного древовидного графа (с англ. – «*tree-structured graph*», «*tree-structured kinematic chain*») [7, 8]. Обобщенное представление подобной модели представлено на рисунке 1 [9].

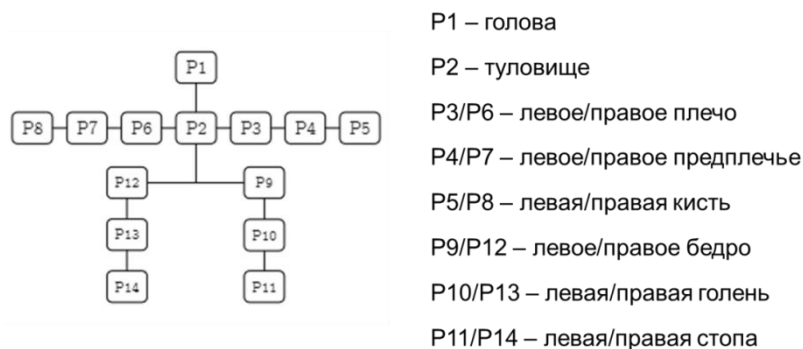


Рисунок 1. Обобщенная модель тела спортсмена в формате упорядоченного древовидного графа [9]

Данный формат модели содержит в себе необходимый и достаточный объем информации, позволяет отразить все ключевые биокинематические звенья и сочленения, а также структурные взаимосвязи между ними.

Подобные модели рассматриваются в пределах некоторого пространства (чаще всего, трехмерного) и данные, характеризующие относительное и абсолютное пространственное положение каждого угла представляют собой проекцию на координатные оси, определяющие ориентацию пространства. Объем и формат данных, получаемых с использованием кинематической модели структуры упорядоченного древовидного графа, различается в зависимости от области применения и решаемой задачи. Наиболее распространенным исходным форматом представление данных, характеризующих двигательные действия человека, является матричное представление. В рамках подобного представления данные подразделяются на отдельные совокупности, каждая из которых соответствует пространственной позе в текущий момент времени (рисунок 2).

$$\begin{array}{c}
 \begin{array}{c} \downarrow \\ n \end{array} \left[\begin{array}{c} \begin{array}{c} \xrightarrow{m} \\ [x_{01}, y_{01}, z_{01}], [x_{02}, y_{02}, z_{02}], \dots [x_{0i}, y_{0i}, z_{0i}] \\ [x_{11}, y_{11}, z_{11}], [x_{12}, y_{12}, z_{12}], \dots [x_{1i}, y_{1i}, z_{1i}] \\ \dots \\ [x_{j1}, y_{j1}, z_{j1}], [x_{j2}, y_{j2}, z_{j2}], \dots [x_{ji}, y_{ji}, z_{ji}] \end{array} \end{array} \right.
 \end{array}$$

Рисунок 2. Матричное представление кинематики двигательных действий человека в 3-х мерном пространстве

Содержание матрицы представляет собой пространственные координаты каждого узла, совокупность которых составляет целостную кинематическую модель. Размер матрицы определяется количеством строк и столбцов ($m \times n$, где m – количество узлов, используемых для описания тела человека в формате кинематической модели; n – количество моментов времени, в течение которых осуществлялось двигательное действие) [10]. Подобное представление является наиболее удобным и распространенным для количественного описания двигательных действий человека, которое представляет собой совокупность последовательных поз в пространстве. Для каждой из этих поз характерно уникальная совокупность пространственных координат каждого узла. Однако главными недостатками данного представления является наличие погрешности, связанной с антропометрическими размера тела спортсмена и отдельных его сегментов, а также малоинформативное содержание данных, особенно при их анализе в масштабе одного момента времени (одна поза). В таком случае получается не полноценная матрица, а вектор-строка.

Нормированная матрица расстояний. Более совершенным и содержательным форматом и структурой данных является нормированная матрица расстояний, которая содержит в себе числовые значения расстояний одних сочленений относительно других (рисунок 3) [11].

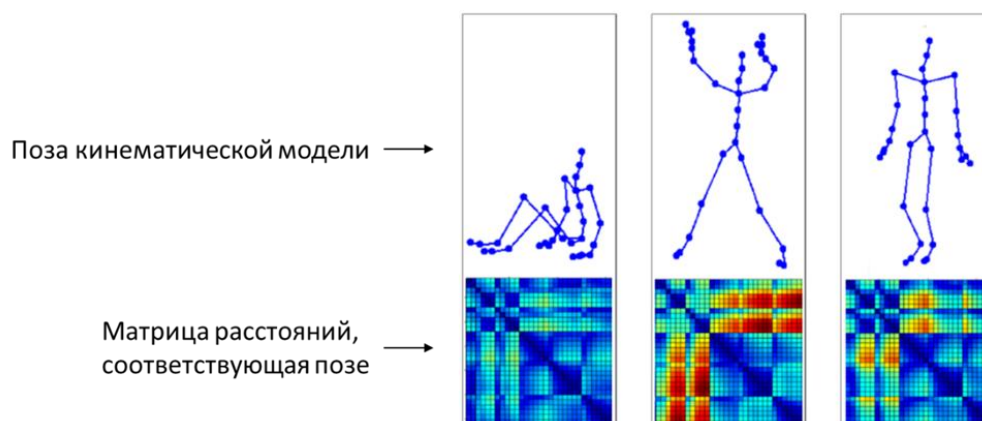


Рисунок 3. Визуализация представления кинематической модели тела человека на основании матрицы расстояний [11]

Наиболее распространенной и фундаментальной метрикой для формирования матрицы расстояний (вне зависимости от размерности пространства) является евклидова метрика, определяющая геометрическое расстояние между точками в n -мерном евклидовом пространстве (1) [12]:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (1)$$

где $d(p, q)$ – геометрическое расстояние между точками p и q в n -мерном пространстве;
 n – размерность пространства.

Однако расчета значений евклидовой метрики недостаточно. Необходимо также осуществить нормализацию данных. Одним из способов нормализации является вычисления отношения длины цепи к пространственному расстоянию (значение евклидовой метрики). Длине цепи соответствует фактическое расстояние между точками (сумма длин звеньев, входящих в цепь). Данное отношение рассчитывается по формуле (2) [13]:

$$dn(p, q) = \frac{d(p, q)}{r(p, q)} \quad (2)$$

где $dn(p, q)$ – нормализованное геометрическое расстояние между точками p и q в n -мерном пространстве;

$r(p, q)$ – фактическая длина цепи.

Нормированная матрица расстояний в анализе и контроле технической подготовленности спортсменов. Нормированная матрица расстояний в масштабе позы позволяет сохранить всю полноту биомеханических признаков, характерных для того или иного двигательного действия. К числу таких признаков относятся прежде всего пространственные, выраженные в величинах суставных углов всех сочленений, используемых в структуре модели. Однако величины суставных углов в матрице не представлены явно, а обозначены через значения расстояния тех или других сочленений относительно друг друга. В частности, на рисунках 4-6 представлены примеры нормированных матриц расстояний для ключевых поз в технике бега на 400 м.

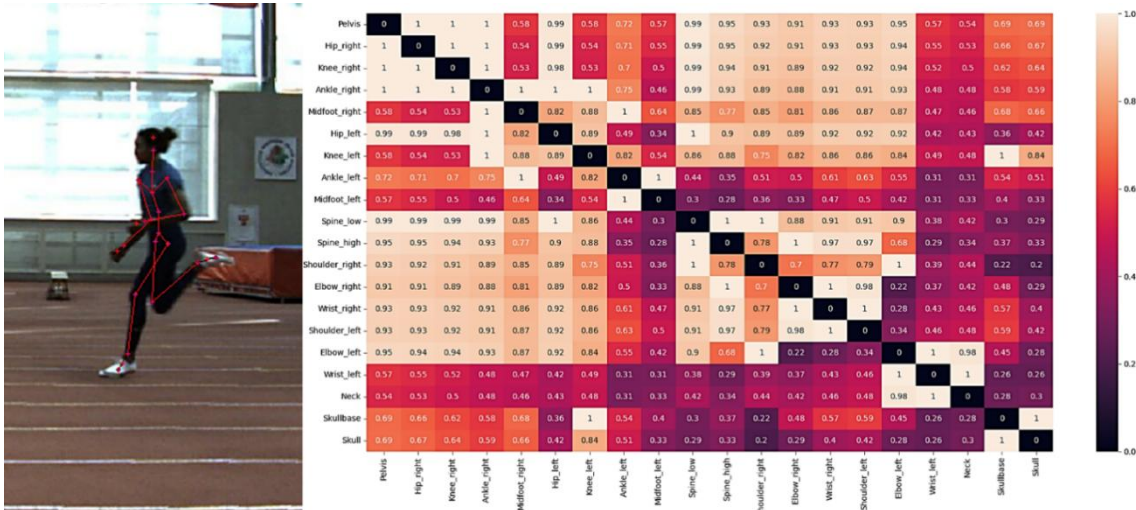


Рисунок 4. Нормированная матрица расстояний для момента контакта спортсмена с поверхностью опоры

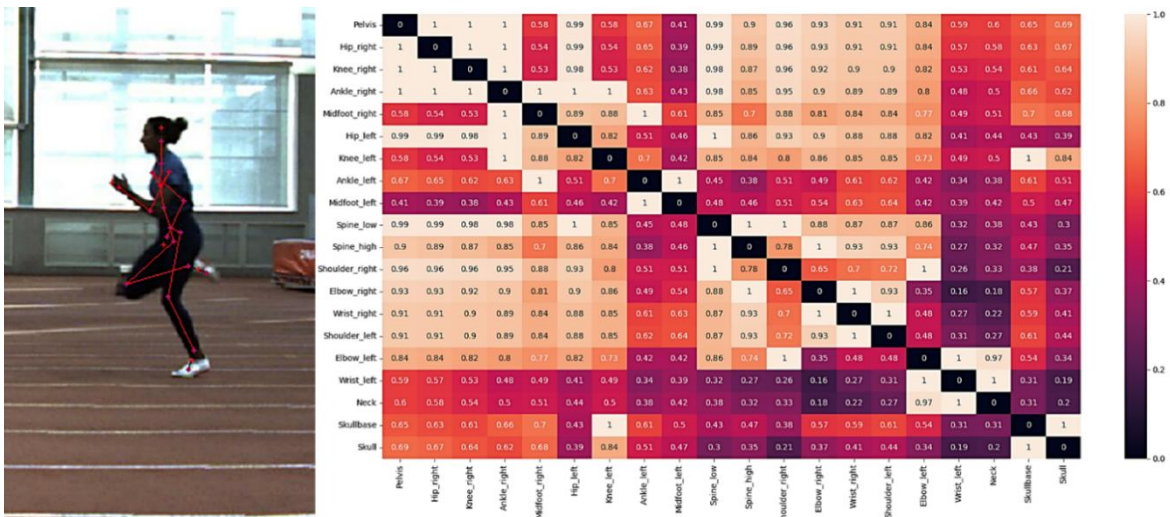


Рисунок 5. Нормированная матрица расстояний для момента амортизации спортсмена

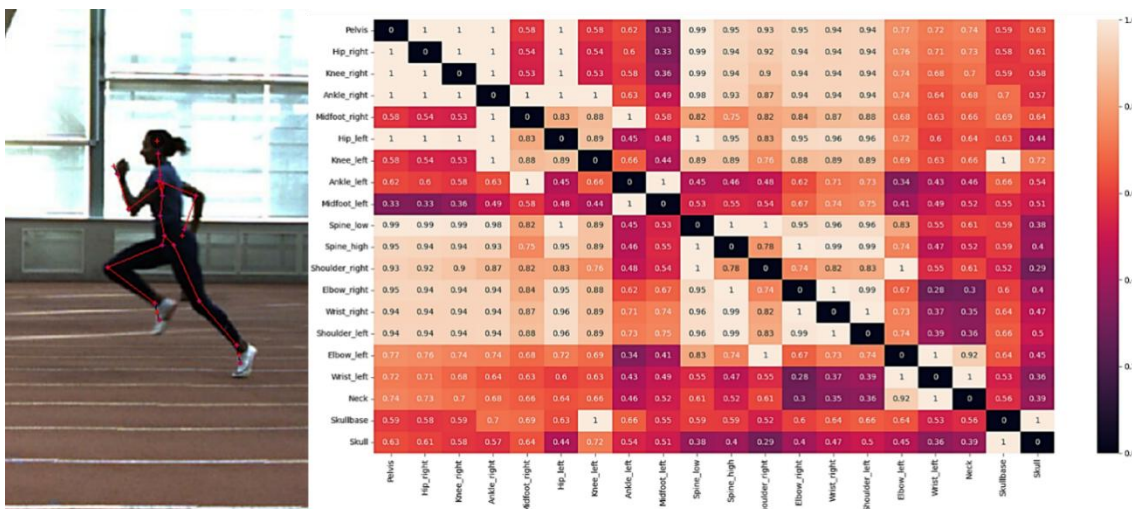


Рисунок 6. Нормированная матрица расстояний для момента отрыва спортсмена от поверхности опоры

На основании данных, содержащихся в матрицах, имеется возможность анализировать технику двигательных действий спортсмена на предмет соответствия модельных характеристикам, сравнивать технику двигательных действий спортсменов друг с другом, контролировать и отслеживать стабильность двигательного стереотипа в зависимости от этапа подготовки. Кроме того, такая структура и формат представления данных в силу своей упорядоченности и структурной стандартизованности, может быть использована для дальнейшей обработки в качестве входных данных в составе различных вычислительных алгоритмов, описывающих принципы работы систем автоматизированного анализа и контроля технической подготовленности спортсменов.

Заключение. Системы захвата движений позволяют получать большой объем полезной и содержательной информации. Главной проблемой и ограничением внедрения таких систем в различные области жизнедеятельности заключается в сложности структурирования и форматирования такой информации. Нормированные матрицы расстояний для отдельных поз позволяют решать множество актуальных прикладных задач в области спортивной биомеханики, поскольку позволяют количественно учесть пространственную сложность двигательных действий, а также сохранить их ключевые биомеханические признаки.

Список литературы

- [1] Bortolini M., Faccio M., Gamberi M., Pilati F. Motion Analysis System (MAS) for production and ergonomics assessment in the manufacturing processes. *Computers & Industrial Engineering*. 2020;139:105485. <https://doi.org/10.1016/j.cie.2018.10.046>
- [2] Menolotto M., Komaris D., Tedesco S., O'Flynn B., Walsh M. Motion Capture Technology in Industrial Applications: A Systematic Review. *Sensors (Basel)*. 2020;20(19):5687. <https://doi.org/10.3390/s20195687>
- [3] Платонов В.Н. Двигательные качества и физическая подготовка спортсменов. К.: Олимп. лит., 2017. – 656 с.
- [4] Sian B., Button C. A Review of Vision-Based Motion Analysis in Sport. *Sports Medicine*. 2008;38(12):1025–1043. <https://doi.org/10.2165/00007256-200838120-00006>
- [5] Шапиро Л., Стокман Дж. Компьютерное зрение. М.: Бином. Лаборатория знаний, 2006. – 752 с.
- [6] Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. М.: «Вильямс», 2004. – 928 с.
- [7] Sedmidubsky J., Elias P., Zezula P. Effective and efficient similarity searching in motion capture data. *Multimedia Tools and Applications*. 2017;77:12073–12094. <https://doi.org/10.1007/s11042-017-4859-7>
- [8] Wang J., Tan S., Zhen X., Xu S., Zheng F., He Z., Shao L. Deep 3D human pose estimation: A review. *Computer Vision and Image Understanding*. 2021;210:103225. <https://doi.org/10.1016/j.cviu.2021.103225>
- [9] Khan M. H., Schneider M., Farid M.S., Grzegorzec M. Detection of Infantile Movement Disorders in Video Data Using Deformable Part-Based Model. *Sensors*. 2017;18(10):3202. <https://doi.org/10.3390/s18103202>
- [10] Parag A., Ketaki A., Prabhakaran B. Robust blind watermarking mechanism for motion data streams. In *Proceedings of the 8th workshop on Multimedia and security (MM&Sec '06)*. Association for Computing Machinery, New York, NY, USA. 2006:230–235. <https://doi.org/10.1145/1161366.1161407>
- [11] Vieira A.W., Schwartz W.R., Campos M., Lewiner T. Distance Matrices as Invariant Features for Classifying MoCap Data. *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, Tsukuba, Japan. 2012:2934–2937.
- [12] Deza M.M., Deza E. *Encyclopedia of Distances*. Springer, 2014. – P 103.
- [13] Bernardin A., Hoyet L., Mucherino A., Gonçalves D., Multon F. Normalized Euclidean distance matrices for human motion retargeting. 2017:1–6. <https://doi.org/10.1145/3136457.3136466>

Авторский вклад

Гусейнов Даниил Истамович – описание принципа формирования матриц расстояний, описание нормированной матрицы расстояний в анализе и контроле технической подготовленности спортсменов, сбор, обработка, анализ и визуализация экспериментальных материалов.

Давыдова Надежда Сергеевна – постановка задачи исследования, описание кинематической модели тела спортсмена, формирование структуры статьи.

KINEMATIC MODEL OF HUMAN MOVEMENTS BASED ON DISTANCE MATRICES

D.I. Guseinov

*PhD student of sports engineering
department of BNTU*

N.S. Davydova

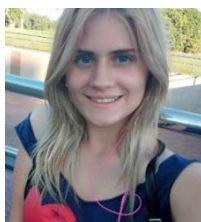
*Associate Professor of the
Academic department of
infocommunication technologies,
PhD of Technical Sciences,
Associate Professor*

Abstract. The article considers an approach to the analysis and control of sports movement techniques using motion capture system data and normalized distance matrices. The kinematic structure of the athlete's body is analyzed, as well as some aspects of the formation of the content of normalized distance matrices. The distance matrix for the key poses of an athlete in the 400 m running technique is given.

Keywords: motion capture systems, matrices, normalization, kinematics, model.

УДК 004.056.5:004.421.22:658.8

УМНЫЕ ДОМА И DATA SCIENCE: ПРЕОБРАЗОВАНИЕ ДАННЫХ В ПРОИЗВОДСТВЕННЫХ И БИЗНЕС СРЕДАХ. ПРЕИМУЩЕСТВА И НЕДОСТАТКИ



А.Д. Каменкова

Бакалавр, инженер по электронным системам,
инженер встраиваемых систем и информационной
безопасности в ООО «Рокет Проджект»
akamend@gmail.com

Каменкова А.Д.

Окончила Белорусский государственный университет информатики и радиоэлектроники по специальности «Программируемые мобильные системы», продолжила обучение в магистратуре по специальности информационные системы и технологии (профилизация: цифровизация в психологии управления человеческими ресурсами). Область научных интересов связана с исследованием программно-аппаратных комплексов систем «умный дом», встроенных систем и микроконтроллеров, сенсорики и датчиков, беспроводных коммуникаций, интернет вещей (IoT), управление энергопотреблением, а также в областях разработки программного обеспечения.

Аннотация. В докладе рассматриваются вопросы применения в промышленности получившего широкое распространение феномена больших данных. Акцентируется внимание на выявлении преимуществ применения *Data Science* в современной промышленности.

Ключевые слова: анализ данных, технологии, промышленность, системы.

Введение. Современный мир данные пронизывают все сферы жизни, особенно в сфере бизнеса и производства. Однако сами по себе данные не обладают ценностью - их нужно анализировать и преобразовывать в полезную информацию, чтобы получить конкурентные преимущества. В этом контексте *Data Science* играет ключевую роль, позволяя компаниям находить скрытые закономерности и паттерны в данных, а также сделать предсказания, которые могут помочь в принятии более обоснованных решений. В данном докладе мы рассмотрим важность преобразования данных *Data Science* и его влияние на производственные и бизнес-преимущества.

Анализ производственных и бизнес преимуществ при использовании Data Science.

1 Анализ данных. *Data Science* обеспечивает компаниям возможность проводить глубокий анализ больших объемов данных. С помощью алгоритмов машинного обучения и статистических методов, *Data Science* позволяет проводить предобработку данных, выявлять выбросы и недостающие значения, а также строить модели, которые могут предсказывать будущие события. Анализ данных помогает выявлять скрытые закономерности и паттерны, которые могут быть использованы для оптимизации производственных процессов и бизнес-стратегий. Например, на основе анализа данных можно определить оптимальные параметры производства, такие как скорость, температура или влажность, которые максимизируют качество продукции.

Анализ данных помогает выявлять потребности и предпочтения клиентов, что может быть использовано для улучшения маркетинговых стратегий и создания персонализированной продукции.

2 Прогнозирование и оптимизация. *Data Science* позволяет строить модели прогнозирования, которые могут помочь компаниям в принятии более точных и обоснованных решений. Например, модели прогнозирования спроса на продукцию могут помочь компаниям планировать производство и складские запасы, избегая лишних затрат или дефицита товаров. Оптимизация процессов, основанная на определенных прогнозируемых данных, также может привести к снижению издержек и улучшению качества продукции. Например, с помощью алгоритмов машинного обучения можно оптимизировать рабочие графики, маршруты доставки или процессы сборки, что позволит снизить время выполнения задач и повысить эффективность производства.

Data Science позволяет компаниям анализировать данные о качестве продукции и выявлять причины возникновения дефектов. На основе этих данных можно принять меры по улучшению производственных процессов или внести изменения в дизайн продукции, чтобы уменьшить количество отказов и повысить удовлетворенность клиентов.

Анализ данных также позволяет компаниям проводить мониторинг качества продукции в реальном времени. Это позволяет своевременно выявлять проблемы и вмешаться до того, как они приведут к серьезным последствиям.

3 Расширение границ. *Data Science* открывает новые возможности для бизнеса и производства, позволяя работать с данными, которые ранее были недоступны или сложны для анализа. Например, анализ данных с *IoT*-устройств (интернета вещей) позволяет получать данные в реальном времени о состоянии оборудования и производственных процессов, что помогает предупреждать аварии и снижать вероятность простоев.

Data Science также позволяет компаниям работать с неструктурированными данными, такими как тексты, изображения или звуки. Анализ таких данных может помочь автоматизировать процессы, улучшить качество контента и обслуживания клиентов [1].

Data Science позволяет компаниям анализировать данные о потребителях и идентифицировать новые сегменты рынка. Анализ данных может помочь определить профили потенциальных клиентов, их потребности и предпочтения, что позволяет компаниям разработать целевые маркетинговые стратегии и привлечь новых клиентов.

4 Принятие обоснованных решений. *Data Science* обеспечивает компаниям аналитическую базу для принятия обоснованных и прогнозируемых решений. На основе анализа данных и построенных моделей компания может прогнозировать результаты различных стратегических решений и выбирать наиболее оптимальные варианты.

Data Science также помогает компаниям проводить *A/B*-тестирование и эксперименты, чтобы определить эффективность новых подходов или технологий. Это позволяет ускорить процесс принятия решений и уменьшить риски ошибок.

Data Science позволяет компаниям получать ценную информацию о рынке, потребительских предпочтениях и трендах. Анализ данных помогает идентифицировать новые возможности для развития бизнеса, выявлять конкурентные преимущества и прогнозировать изменения в потребительском поведении.

5 Оптимизация процессов. С помощью *Data Science* компании могут анализировать данные о своей производственной деятельности и выявлять узкие места, неэффективные процессы и возможности для оптимизации. Анализ данных может помочь компаниям улучшить планирование производства, снизить издержки, повысить эффективность использования ресурсов и улучшить общую производительность компании.

6 Улучшение качества продукции. Технология *Data Science* позволяет компаниям анализировать данные о качестве продукции и идентифицировать проблемные области. Анализ данных может помочь компаниям выявить причины дефектов и неполадок в производственных процессах, что позволяет им принимать меры для улучшения качества продукции. Это способствует повышению удовлетворенности клиентов, снижению риска возвратов и репутационным преимуществам.

Следует отметить, что использование *Data Science* имеет множество преимуществ для компаний в разных отраслях. Он помогает компаниям стать более эффективными, адаптивными и успешными в современном бизнес-мире, где данные играют все более важную роль в принятии обоснованных решений и достижении конкурентного преимущества. Поэтому инвестиции в *Data Science* становятся неотъемлемой частью стратегии развития компании.

Ознакомившись с преимуществами *Data Science*, попробуем применить их для оптимизации производства и получения бизнес преимуществ в производстве системы умного дома.

Оценим возможность интегрирования *Data Science* в систему умного дома, а также компании по производству данных систем [2]:

1 Оценка рынка систем «Умный дом».

Использование *Big data* позволяет компании при минимизации затрат получить максимально репрезентативный отчет о рынке данных систем, что позволит выявить наиболее популярные и востребованные системы, а также сформировать требования к разрабатываемой системе.

2 Оценка самых продаваемых систем.

Данный этап позволяет на основе алгоритмов и машинного обучения выявить наиболее продаваемые системы. Важно отметить что Представленные данные будут отражать не только объективные и часто упоминаемые факторы (такие как используемые компоненты системы), но и субъективные факторы (отзывы покупателей).

3 Формирование основных технических требований к системе при помощи алгоритмов *Data Science*.

Использование *Data Science* позволит выбрать наиболее часто используемые компоненты в системах умного дома, а также подобрать компоненты с наилучшим соотношением цена/качество.

4 Использование *Data Science* для формирования производственных мощностей и складских запасов.

5 Использование данных для оптимизации и персонализации системы умного дома.

Как можно заметить, использование *Data Science* позволяет оптимизировать производство, снизить издержки и подобрать наиболее подходящую бизнес-стратегию для рассматриваемого рынка. Однако, следует более подробно остановиться на использовании данных для оптимизации систем умного дома.

Данное решение не требует мощного аппаратного комплекса, однако может быть реализовано программно, что добавит больше возможностей системы умного дома, при незначительном повышении цены данной системы [3]. Алгоритм разработки для производства данной системы необходимо использовать следующий:

1 Предварительная обработка данных. Важной частью *Data Science* является предварительная обработка данных. Это включает в себя удаление шума, обработку пропущенных значений, масштабирование данных и другие техники, которые позволяют получить чистый и точный набор данных для анализа. В случае умных домов, предварительная обработка данных может включать в себя фильтрацию аномальных показателей потребления энергии или других параметров, а также приведение временных рядов к удобному формату для анализа.

2 Анализ данных. При предварительной обработке данных следует анализ, который может включать в себя использование различных методов машинного обучения, статистики и исследовательского анализа данных. В случае умных домов, анализ может помочь выявить паттерны потребления энергии, предсказать будущее потребление, оптимизировать энергетические режимы и т.д. Дополнительно, анализ данных может помочь выявлять

аномалии в работе устройств и оперативно реагировать на них, минимизируя время простоя и улучшая обслуживание.

3 Принятие решений на основе данных. После анализа данных становится возможным принятие обоснованных решений на основе полученных выводов. Например, анализ показателей потребления энергии в умном доме может помочь в принятии решений о замене устаревших энергосберегающих устройств или оптимизации энергетического расписания. Это приводит к снижению расходов на энергию, улучшению комфортности проживания и, в итоге, к повышению конкурентоспособности жилья.

4 Бизнес-преимущества. Применение *Data Science* в умных домах может привести к ряду бизнес-преимуществ. Во-первых, оптимизация потребления энергии может снизить затраты на энергию для владельцев домов и жильцов. Во-вторых, анализ данных позволяет предсказать возможные поломки и проблемы с оборудованием, что позволяет оперативно проводить предупредительное обслуживание и сокращать расходы на ремонт. Кроме того, данные о поведении пользователей могут быть использованы для настройки персонализированных услуг, предлагаемых владельцам умных домов, что повышает их удовлетворенность и лояльность.

5 Предсказание и предупреждения. *Data Science* может помочь предсказывать и предупреждать о возможных проблемах и аварийных ситуациях в умном доме. Анализ данных об использовании электроприборов и датчиков безопасности позволяет выявить аномалии и предсказать возможные поломки или неполадки. Это позволяет оперативно принять меры по предотвращению аварийных ситуаций, а также планировать техническое обслуживание и ремонтные работы.

Преимущества Использование *Data Science* в умных домах.

1 Энергоэффективность: Анализ данных потребления энергии позволяет эффективно управлять энергией в умном доме, оптимизировать тарифы и снизить расходы на электроэнергию.

2 Комфорт и удобство: Алгоритмы управления системами умных домов, разработанные на основе анализа данных, позволяют создать оптимальные условия комфорта и удобства для жильцов.

3 Безопасность: Анализ данных от датчиков безопасности позволяет выявлять потенциально опасные ситуации и предупреждать о них, обеспечивая высокий уровень безопасности в умном доме.

4 Оптимизация затрат: Анализ данных помогает выявить потенциальные источники экономии и оптимизировать затраты на энергию, обслуживание и ремонт в умном доме.

5 Автоматизация: *Data Science* позволяет автоматизировать управление системами в умном доме, что упрощает повседневную жизнь жильцов и освобождает время для других задач.

6 Персонализация: Анализ данных позволяет создавать персонализированные рекомендации и настройки для каждого жильца в умном доме. Например, на основе данных о предпочтениях и привычках можно создать индивидуальные профили для управления освещением, температурой и другими параметрами в доме, чтобы каждый член семьи чувствовал себя комфортно.

7 Домовая безопасность: *Data Science* позволяет улучшить системы домашней безопасности в умном доме. Компьютерное зрение и алгоритмы машинного обучения могут обнаруживать подозрительную активность или незапланированные вторжения на территорию дома. Это может включать распознавание лиц и определение подозрительных поведенческих моделей для предупреждения о возможных угрозах.

8 Улучшение жизненного цикла устройств: *Data Science* может помочь в улучшении жизненного цикла устройств в умном доме. Анализ данных о работе и обслуживании устройств позволяет определить оптимальные временные интервалы для технического

обслуживания и замены частей. Это позволяет сократить затраты на ремонт и увеличить эффективность работы устройств.

9 Управление отходами и экологическая устойчивость: *Data Science* может быть использована для управления отходами и повышения экологической устойчивости в умном доме. Анализ данных о потреблении и утилизации материалов позволяет оптимизировать процессы утилизации и переработки отходов. Также можно разрабатывать и применять модели и алгоритмы для управления энергоресурсами, снижая негативное воздействие на окружающую среду.

На основании этих факторов можно сделать вывод, что преобразование данных *Data Science* является одним из ключевых факторов для достижения производственных и бизнес-преимуществ. *Data Science* является неотъемлемой частью современного бизнеса и может существенно повысить его конкурентоспособность. Например, с помощью анализа данных компания может оптимизировать производственные процессы, сократив издержки и увеличивая эффективность. Это может быть достигнуто путем определения оптимальной структуры использования ресурсов, оптимизации маршрутов доставки или сокращения времени настройки оборудования. В то же время, технологии *Data Science* позволяют повышать качество продукта и открывают новые возможности для его использования. Преобразование данных с помощью *Data Science* позволяет оптимизировать потребление энергии, предсказывать поломки и проблемы оборудования, предоставлять персонализированные услуги и многое другое. Применение *Data Science* в области умных домов является эффективным инструментом для улучшения качества жизни, экономии ресурсов и повышения конкурентоспособности на рынке недвижимости.

Список литературы

- [1] Позднякова А.О. Big Data, Data Science и Machine Learning, как новейшие тренды современности / А.О. Позднякова // Аллея науки. – 2022. – Т. 1, № 5(68). – С. 10-16. – EDN AWCВAC.
[2] Lee R. Big Data, Cloud Computing, and Data Science Engineering. – Cham: Springer. – 2020. – 214 p.
[3] Веретенников А. В. Big Data: анализ больших данных сегодня. – 2017. – № 32 (166). – С. 9-12. Машинное обучение. [Электронный ресурс]. URL: <https://blog.skillfactory.ru/glossary/mashinnoe-obuchenie/> (дата обращения: 05.01.2024).

Авторский вклад

Каменкова Анна Дмитриевна – постановка задачи и исследование алгоритма проектирования программно-аппаратных систем «умного дома», анализ применения Data Science в современной промышленности, выделение ключевых преимуществ и недостатков этого подхода, с учетом важности для современных производственных и бизнес-сред, формирование структуры статьи

SMART HOMES AND DATA SCIENCE: DATA TRANSFORMATION IN INDUSTRIAL AND BUSINESS ENVIRONMENTS. ADVANTAGES AND DISADVANTAGES

A.D. Kamenkova

*Bachelor's degree holder, Electronic Systems Engineer,
Embedded Systems Engineer, and Information Security
Engineer at Rocket Project LLC.*

Abstract. The report examines the application of the widely prevalent phenomenon of big data in industry. Emphasis is placed on identifying the advantages of applying Data Science in modern industry.

Keywords. data analysis, technologies, industry, systems.

УДК 004.056.53

МЕТОДЫ И СПЕЦИФИКА ЗАЩИТЫ КОМПЬЮТЕРНЫХ СЕТЕЙ ПРИ РАБОТЕ С BIG DATA



Е.А. Лещенко

Инженер-программист
отдела сетевых технологий
Центра
информатизации и
инновационных разработок
БГУИР, ассистент кафедры
информатики, магистрант
кафедры ПИКС
e.leshchenko@bsuir.by



Е.В. Бегляк

Инженер-программист I
категории отдела сетевых
технологий Центра
информатизации и
инновационных разработок
БГУИР, ассистент кафедры
ВМиП, магистрант кафедры
ПИКС
katarina@bsuir.by



В.А. Марцинкевич

Ведущий инженер-
электроник отдела сетевых
технологий Центра
информатизации и
инновационных разработок
БГУИР, ассистент кафедры
ЭВМ, магистрант кафедры
ИКТ
vlad@bsuir.by

Е.А. Лещенко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Инженер-программист отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики БГУИР, магистрант кафедры проектирования информационно-компьютерных систем БГУИР.

Е.В. Бегляк

Окончила Белорусский государственный университет информатики и радиоэлектроники. Инженер-программист отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры вычислительных методов и программирования БГУИР, магистрант кафедры проектирования информационно-компьютерных систем БГУИР.

В.А. Марцинкевич

Окончил Минский радиотехнический институт. Ведущий инженер-электроник отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры электронных вычислительных машин БГУИР, магистрант кафедры инфокоммуникационных технологий БГУИР.

Аннотация. В данной работе рассмотрены методы и принципы защиты компьютерных сетей, в которых выполняется передача *Big Data*.

Показано, что необходим комплексный подход из различных методов безопасной передачи информации для защиты компьютерных сетей при работе с большими данными.

Ключевые слова: защита компьютерной сети, шифрование, передача больших данных

Введение. В современном информационном обществе компьютерные сети являются неотъемлемой частью бизнес-процессов и повседневной жизни многих людей. Однако, с ростом зависимости от компьютерных сетей, возрастает объем данных, генерируемых, собираемых анализируемых вычислительными системами и вопросы безопасности становятся все более актуальными. Особенно важно обеспечить защиту при работе с большими данными, так как такие сети обычно содержат большое количество

конфиденциальной информации. Злоумышленники могут использовать различные методы для вторжения в компьютерные сети и получения несанкционированного доступа к конфиденциальной информации.

Безопасность больших данных. Сегодня почти каждая организация рассматривает возможность внедрения больших данных, потому что они видят их потенциал и пытаются его использовать. Независимо от размера организации, каждый пытается защитить свои данные.

Согласно отчету IBM и Института Понемона за 2023 год, средняя стоимость утечки данных в 2023 году достигнет 4,45 миллиона долларов США, увеличившись на 2% по сравнению с 2022 годом (4,35 миллиона долларов США).

Обеспечение безопасности больших данных затруднено по нескольким причинам.

Некоторые из них упомянуты ниже:

– В режиме реального времени поступает множество данных из разных источников с разными потребностями в защите.

– Существует несколько типов данных, объединенных вместе.

– Доступ к данным получают множество разных пользователей с различными аналитическими требованиями.

– Быстро развивающиеся инструменты, финансируемые сообществом открытого исходного кода.

– Автоматическая репликация данных между несколькими узлами [1].

Методы защиты компьютерных сетей. Методы защиты компьютерной сети при работе с большими данными включают в себя несколько аспектов. Во-первых, необходимо обеспечить защиту самой сети от внешних угроз. Во-вторых, необходимо обеспечить защиту данных, хранящихся и передаваемых по сети. Большие данные часто содержат конфиденциальную информацию, и их утечка может привести к серьезным последствиям. Для защиты данных применяются различные методы шифрования, аутентификации и авторизации.

Защита сети от внешних угроз включает в себя:

1 Использование брандмауэра – это программное или аппаратное устройство, которое контролирует и фильтрует трафик, проходящий через сеть. Он позволяет установить правила доступа к сети и блокировать подозрительный или вредоносный трафик. Брандмауэр также может обнаруживать и предотвращать атаки на сеть, такие как *DDoS*-атаки или попытки взлома. Брандмауэр также может предоставлять функции аутентификации и шифрования данных, а также возможность мониторинга и журналирования сетевой активности.

2 Использование виртуальных частных сетей (*VPN*) – создание защищенного канала связи между удаленными участниками сети.

3 Контроль доступа – установление политик и правил, определяющих, кто имеет доступ к данным и какие операции он может выполнять.

4 Шифрование данных – преобразование информации в неразборчивый вид для посторонних лиц.

5 Резервное копирование – создание резервных копий данных для их восстановления в случае катастрофы или взлома.

Использование антивирусного программного обеспечения – для обнаружения и удаления вредоносных программ, таких как вирусы, трояны, шпионское ПО и другие угрозы безопасности. Она может быть установлена на отдельные компьютеры в сети или на центральный сервер для защиты всех устройств в сети. Антивирусная программа сканирует файлы и систему на наличие вредоносных программ, блокирует их действия и предупреждает пользователя о потенциальных угрозах. Она также может обновляться

регулярно для получения новых определений вирусов и обновлений программы, чтобы эффективно бороться с новыми угрозами.

Принципы защиты компьютерной сети при работе с большими данными.

Одним из принципов защиты компьютерной сети является принцип делегирования полномочий. Делегирование полномочий в компьютерной сети – это процесс передачи определенных прав и обязанностей от одного пользователя или группы пользователей другому пользователю или группе пользователей. Это позволяет распределить ответственность и управление в сети, обеспечить эффективное выполнение задач, а также это ограничивает возможности злоумышленников в случае взлома аккаунта или устройства. Делегирование полномочий может включать следующие аспекты:

1 Права доступа. Администратор сети может делегировать определенные права доступа к файлам, папкам, приложениям или другим ресурсам сети. Например, администратор может предоставить право на чтение, запись или выполнение определенных файлов или папок пользователю или группе пользователей.

2 Управление учетными записями. Администратор сети может делегировать управление учетными записями, включая создание, удаление и изменение учетных записей пользователей или групп.

3 Управление политиками безопасности. Администратор сети может делегировать управление политиками безопасности, включая настройку правил брандмауэра, шифрование данных, установку антивирусного программного обеспечения и т. д.

4 Мониторинг и аудит. Администратор сети может делегировать задачи мониторинга и аудита определенным пользователям или группам.

Регулярная установка обновлений и исправлений программного обеспечения для закрытия уязвимостей также необходима при работе с большими данными. Уязвимости могут быть использованы злоумышленниками для взлома системы, поэтому регулярное обновление является важным аспектом безопасности. Приложения, работающие в компьютерной сети, должны быть защищены от уязвимостей и возможных атак. Для этого можно использовать патчи безопасности, аудит кода и использование безопасных разработочных практик.

Согласно принципу непрерывности работы компьютерной сети, система должна быть способна продолжать свою работу даже в случае возникновения сбоев или атак. В случае возникновения непредвиденных событий или катастрофических ситуаций этот метод позволяет минимизировать потери и простои в работе. В компьютерной сети он включает в себя меры по защите и восстановлению данных, приложений и инфраструктуры, чтобы минимизировать потери и простои в работе.

Принцип обучения пользователя основам безопасности состоит в том, чтобы обучить пользователей правильным практикам и процедурам, которые помогут им избежать угроз безопасности и защитить компьютерную сеть от потенциальных атак. Они должны понимать, какие действия могут представлять угрозу и какие меры предосторожности следует принять. Основные аспекты этого подхода включают в себя осведомленность о рисках; обучение пользователя основам безопасности, включая правила использования паролей, распознавание подозрительных электронных писем и ссылок и т. д.; постоянное обновление знаний; обратная связь и наказание за нарушение правил безопасности, чтобы создать ответственность и мотивацию для соблюдения правил; регулярные проверки и аудиты.

Шифрование больших данных. Как было сказано выше, необходимо обеспечить защиту данных, хранящихся и передаваемых по сети. Большие данные часто содержат конфиденциальную информацию, и их утечка может привести к серьезным последствиям. Для защиты данных применяются различные методы шифрования, аутентификации и авторизации.

С помощью ключей шифрования компьютерный алгоритм преобразует текстовые символы в непонятную форму, гарантируя, что только авторизованные лица, обладающие необходимыми ключами, смогут разблокировать и получить доступ к контенту. В определенной степени важно обеспечить безопасность различных форм данных, включая файлы, базы данных и сообщения электронной почты.

Одним из основных подходов к шифрованию больших данных является использование асимметричного шифрования. Этот метод основан на использовании пары ключей: открытого и закрытого. Открытый ключ используется для шифрования данных, а закрытый ключ – для их расшифровки. Такой подход позволяет безопасно передавать данные, так как открытый ключ может быть распространен публично, в то время как закрытый ключ хранится в секрете у владельца данных.

Для шифрования больших данных также широко применяются алгоритмы блочного шифрования, такие как *AES (Advanced Encryption Standard)* и *Blowfish*. Эти алгоритмы разбивают данные на блоки фиксированного размера и применяют к каждому блоку операции шифрования. Блочные алгоритмы обеспечивают высокую степень безопасности и эффективности при шифровании больших объемов данных [2].

Кроме того, при шифровании больших данных важным аспектом является управление ключами. Ключи шифрования должны быть хранены в безопасном месте и доступны только авторизованным пользователям. Для обеспечения безопасности ключей можно использовать различные методы, включая аппаратное шифрование и многофакторную аутентификацию. Также необходимо регулярное обновление шифровальных ключей для предотвращения возможности их взлома.

Важно отметить, что шифрование больших данных может быть ресурсоемким процессом, особенно при использовании сильных алгоритмов шифрования. Поэтому для эффективного шифрования больших объемов данных могут применяться различные оптимизации, такие как параллельное шифрование и использование специализированных аппаратных решений.

Эффективность шифрования проистекает из его способности делать данные нерасшифрованными, даже когда злоумышленники перехватывают пакеты данных или получают доступ к конфиденциальным файлам.

Шифрование каналов связи является одним из наиболее важных аспектов современной информационной безопасности.

Основной целью шифрования каналов связи является обеспечение конфиденциальности передаваемой информации. Путем применения криптографических алгоритмов и протоколов (как пример, использование *SSL/TLS* протокола для защиты коммуникаций между клиентом и сервером) можно предотвратить несанкционированный доступ к данным и гарантировать, что только сами участники общения могут расшифровать сообщения [3].

Кроме конфиденциальности, шифрование каналов связи также обеспечивает целостность передаваемой информации. Целостность означает, что данные не могут быть изменены в процессе передачи без обнаружения этого факта.

Заключение. Таким образом, защита компьютерной сети при работе с большими данными является сложным и многогранным процессом. Она требует применения различных методов и технологий, а также постоянного мониторинга и анализа угроз. Только комплексный подход к защите может обеспечить надежность и безопасность работы с большими данными.

Список литературы

[1] IBM Security опубликовала отчет о стоимости утечки данных за 2023 год: [Электронный ресурс]. URL: <https://www.codetd.com/ru/article/16598851>. (Дата обращения: 10.02.2024).

[2] Современные зарубежные шифры: [Электронный ресурс]. URL: <https://хакер.ру/2016/06/30/crypto-part4/>. (Дата обращения: 12.02.2024).

[3] Big Data – Data Encryption in Big Data: [Электронный ресурс]. URL: <https://www.encryptionconsulting.com/data-protection-in-big-data-using-encryption/>. (Дата обращения: 13.02.2024).

Авторский вклад

Авторы внесли равноценный вклад

METHODS AND SPECIFICS OF COMPUTER NETWORK PROTECTION WHEN WORKING WITH BIG DATA

E.A. Leshchenko

Software Engineer of the Network Technology Department of the Center of Informatization and Innovation Elaborations, Assistant of the Department of Computer Science of BSUIR, Master's student of the Department of ICSD of BSUIR

E.V. Beglyak

Software Engineer of the 1st category of the Network Technology Department of the Center of Informatization and Innovation Elaborations, Assistant of the Department of Computational Methods and Programming of BSUIR, Master's student of the Department of ICSD of BSUIR

V.A. Martsinkevich

Leading Electronics Engineer of the Network Technology Department of the Center of Informatization and Innovation Elaborations, Assistant of the Department of Electronic Computing Machines of BSUIR, Master's student of the Department of Information and Communication Technologies of BSUIR

Abstract. This article describes the methods and principles of protecting computer networks in which Big Data is transmitted.

It is shown that an integrated approach is needed from various methods of secure information transmission to protect computer networks when working with big data.

Keywords: computer network protection, encryption, big data transmission

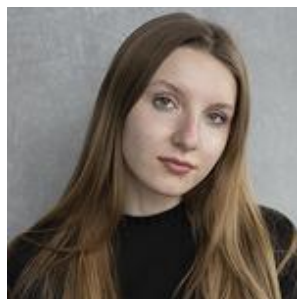
UDC 004.021:004.75

RESEARCH OF LARGE LANGUAGE MODELS FOR TEXT GENERATION AND THEIR PRACTICAL APPLICATION



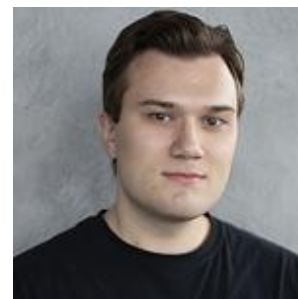
A.N. Markov

Senior Lecturer of the
Department of Informatics,
Deputy Head of the Center for
Informatization and Innovative
Developments of the Belarusian
State University of Informatics
and Radioelectronics
a.n.markov@bsuir.by



M.M. Zyryanova

Student of the Faculty of
Computer Systems and
Networks of the Belarusian
State University of
Informatics and
Radioelectronics.
z.y.r.y.a.n.o.v.a@mail.ru



A.E. Asadchy

Student of the Faculty of
Computer Systems and Networks
of the Belarusian State University
of Informatics and
Radioelectronics.
aleh.asadchy@gmail.com

A. Markov

Graduated from Belarusian State University of Informatics and Radioelectronics. The field of scientific activity includes such areas as computing systems, cloud computing, distributed computing systems, load balancing of computer systems.

M. Zyryanova

A fourth-year student of the Belarusian State University of Informatics and Radioelectronics. The scope of scientific activity includes such areas as large language models and their additional training, as well as effective implementation.

A. Asadchy

A fourth-year student of the Belarusian State University of Informatics and Radioelectronics. The scope of scientific activity includes such areas as large language models and their additional training, as well as effective implementation.

Abstract. This research is aimed at the study of large language models designed for translation and text generation. In the course of the study, transformers of the Text2TextGeneration type were considered, their comparative analysis and testing of various situations were carried out. Large language models were also implemented into an existing software product.

Keywords: large language models, transformers, Text2TextGeneration, fine-tuning, hyperparameters, neural network, Big Data.

Introduction. The field of modern information technologies is going through a period of rapid development, which the average user was able to notice a few years ago, with the advent of widespread use of large language models. In recent years, this area has begun to develop more and more. Large language models based on deep learning and the mechanism of self-attention are able to generate texts of various lengths and contents, which can sometimes be difficult to distinguish from those written by a person. In addition, such models have already become assistants to humans, in many respects to programmers, in their work.

In the context of this research, the aim is to investigate such large language models and their potential implementation in various fields and solutions, including automatic translation, text generation and rewriting, grammar checking, etc.

Research methodologies. The key tasks are translating text from American English to British, translating from English to German, improving the quality of the text, and checking grammar and spelling. In the course of studying various platforms, it was decided to use the Huggingface platform, which provides wide access to various kinds of models and their images absolutely free of charge.

The following criteria were developed to evaluate and select large language models:

- what was taught on what it was taught;
- dataset volume;
- perplexity (how well the model predicts the text);
- relevance (relevance, i.e. the correspondence of the answer to what the person asked);
- coherence;
- overall quality;
- accuracy (accuracy of the data directly reported, i.e. the truthfulness of the data in relation to itself without taking into account the context);
- latency (delay in response from the network, usually delay to the first character);
- context;
- vulnerability of language models;
- Rapid deployment and testing of models.

We will consider and conduct a comparative analysis of large language models in the field of translation from American English to British. Two models were chosen: t5-base-us-to-uk-english and autonlp-US-to-UK-604417040.

Table 1. Comparative Analysis of Models for Translation from American English to British

Criterion	t5-base-us-to-uk-english	autonlp-US-to-UK-604417040
Dataset	249525 American and British sentences	604417040 sentences with American and British spelling
Dataset Volume	249525 offers	604417040 offers
Perplexity	5.2	4.8
Relevance	Average	High
Coherence	High	High
Overall quality	High	High
Accuracy	High	High
Latency	100-200 ms	50-100 ms
Context	Takes context into account	Takes context into account
Vulnerability of language models	Susceptible to bias, can generate incorrect information	Susceptible to bias, can generate incorrect information
Rapid deployment and testing	Fast	Fast

Both models have their advantages and disadvantages. The autonlp-US-to-UK-604417040 model has lower perplexity, higher relevance, and lower latency than t5-base-us-to-uk-english. However, /t5-base-us-to-uk-english has a higher coherence. In this case, it makes sense to have a

high consistency model for related translation of sentences rather than individual words, so in this case the t5-base-us-to-uk-english model was chosen.

We will consider and conduct a comparative analysis of large language models in the field of translation from English to German. Models such as small100, wmt19-en-de and opus-mt-en-de were chosen.

Table 2. Comparative Analysis of Models for Translation from English to German

Criterion	small100	wmt19-en-de	Opus-MT-en-DE
Dataset	100 languages (small sets)	WMT'19 (English-German)	OPUS (multilingual)
Dataset Volume	1.5M offers	4.5M offers	65M offers
Perplexity	45.0	27.0	21.0
Relevance	Low	Average	High
Coherence	Average	High	Very high
Overall quality	Low	Average	High
Accuracy	Low	Average	High
Latency	Low	Average	High
Context	Limited	Average	High
Vulnerability of language models	High	Average	Low
Rapid deployment and testing	High	Average	Low

We can conclude that small100 is the fastest and simplest model, but it has low accuracy and quality. At the same time, wmt19-en-de is a model with higher quality than small100, but it is less versatile. In contrast to the above-mentioned models, opus-mt-en-de is the most accurate and high-quality model, but it is the slowest and most difficult to deploy.

Despite the low speed of the opus-mt-en-de, its precision makes it ideal for the task at hand.

We will consider and conduct a comparative analysis of large language models in the field of translation from English to German. Models such as t5_Grammar_Checker, t5-Base-Grammar-Correction, BERT-Grammar-Checker were chosen.

Table 3. Benchmarking Models for Grammar Checking

Criterion	T5_Grammar_Checker	T5-Base-Grammar-Correction	BERT Grammar Checker
What they taught	Stack Overflow, Reddit, Wikipedia	Stack Overflow, Reddit	Google Books
Dataset Volume	100M+ words	10M+ words	100B+ words
Perplexity	10.2	12.5	8.9
Relevance	High	Average	Low
Coherence	High	Average	Low
Overall Quality	High	Average	Low
Accuracy	High	Average	Low

End of table 3

Criterion	T5_Grammar_Checker	T5-Base-Grammar-Correction	BERT Grammar Checker
Latency	200-300 ms	100-200 ms	50-100 ms
Context	Takes context into account	Takes context into account	Doesn't take context into account
Vulnerability of language models	Vulnerable to bias, noise, attacks	Vulnerable to bias, noise, attacks	Vulnerable to bias, noise, attacks
Rapid deployment and testing	Average	High	High

T5_Grammar_Checker is the best model out of the three. It has high perplexity, relevance, coherence, overall quality, accuracy and latency. T5-Base-Grammar-Correction is a faster model, but it is less accurate and relevant. BERT-Grammar-Checker is the fastest model, but it is the least accurate, relevant, and consistent. The chatgpt_paraphraser_on_T5_base model was chosen as a model for improving the quality of the text due to the good result during the prompting testing.

The latter was the selection of a model to improve the text. There was a slightly different approach to the choice, since we evaluated the transformed text not as a set of some parameters, but as its content and semantic value. After making comparisons on certain prompts, it was decided to choose the first one from the two models humanin/chatgpt_paraphraser_on_T5_base and ramsrigouthamg/t5_sentence_paraphraser due to the better output text.

Architecture and training of large language models. One of the important parts of the development of large language models was the invention of new neural network architectures. In particular, a significant breakthrough in this area was the mechanism of transformers, which uses self-attention.

The transformers model has a significant difference from the models that preceded it: it efficiently processes text without using either recurrent or convolutional layers, which in turn allows for parallel learning and directly increases the speed of learning.

The architecture of the transformer can be represented in the form of the following layers:

- Embeddings Layers. Converts each word in the input sequence into a vector representation. Word vectors and positions can be combined to convey information about the sequence of words and their relative positions.

- Multi-Head Self-Attention Layers. These layers draw the model's attention to different parts of the input sequence. And the Multi-Head mechanism allows the model to look at the input sequence from different perspectives, so the model better understands the context.

- Feed-Forward NN Nonlinear activation functions for processing data from the self-attention layer.

- Normalization and Residual Connection Layers. They are used to improve the stability of learning and prevent gradient attenuation [1].

The architecture of the transformer is shown in Figure 1.

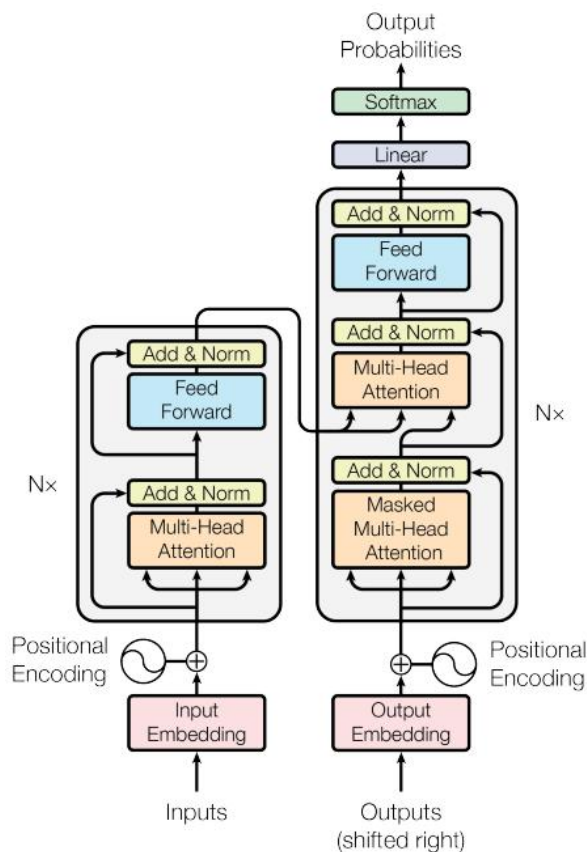


Figure 1. Transformer Architecture

The principle of operation is as follows: the encoder receives words as input, converts them into embeddings by running each word through layers of fully-connected layers, some through shortcut connections, and most importantly through Multi-head attention, i.e. through several attention mechanisms in parallel. Thanks to this mechanism, the model pays attention to the positions of words in a sentence if they are important. Next comes the decoder: it runs one word at a time, by receiving the last word as input and giving the next one. Here again, the self-attention mechanism is used to work with words and their positions, but here we use words that have already been encoding. At the end, there is a regular softmax, which already gives out the probabilities of words. During decoding, each next word interacts with the previous ones and with the encoder vectors [2].

The main advantage of such models is the ability to refer to any word, regardless of the length of the context, as well as the possibility of parallel learning.

Speaking about the learning process, there are several stages:

- Pre-training. Here, the model is trained on a large amount of text data without a specific target. The model simply learns general knowledge about the language and its patterns.
- Fine-tuning. Once a model has been trained on general data, it's time to train it for a specific task where it will be used, such as translation or text generation.
- Hyperparameter tuning is also an important step in training the model.

Use of large language models. Due to their efficiency and ability to work with large long, large language models are applicable in different fields.

One of the most popular areas is text generation itself. This feature is often used by copywriters, as it is well suited for generating articles, reviews, news, or even literary works. This function is highly dependent on hyperparameters, and if you set them up correctly, you can

actually get almost literary works. This feature can also be used to generate content for marketing purposes, advertising slogans, product descriptions, and more.

Automatic translation. Models of this type are successfully used in systems of this type. They are capable of generating high-quality translations between different languages without additional customization in a specific language pair. The best example would be YandexTranslate, DeepL, which openly states the use of transformers architecture within its models.

Summarization of text. Thanks to the model's ability to capture context even in large amounts of text, llms can be used to automatically summarize long texts, and they can also generate concise and informative summaries of articles of various genres, and in general, almost any document.

One of the most popular applications of such models is dialogue systems and chatbots. Due to the model's ability to generate natural and understandable answers to user questions, such systems can be used to communicate with users in their native language. They can be embedded in a huge number of websites, applications, and services. The most popular models at the moment are ChatGPT, Bard. Every day, a huge number of people use it for a variety of purposes: from writing scientific articles and using it for all sorts of household trifles to helping IT specialists in their work.

Thus, large language models are not just a theoretically promising direction of development, but literally practically successful. At the moment, almost every large company wants to have its own model with completely different areas of application, although usually for internal needs, but this shows that this type of model is successful and, moreover, necessary [3].

Implementation of large language models. A topic of direct interest to us was the introduction of large language solutions into ready-made projects in order to increase their efficiency and convenience. Such a project was a CV generation service, which was supplemented by another service containing Transformers.

The main advantage of the project is the use of ready-made transformers models trained on big data and already familiar with the language with which they will work. Using already trained models allows you to save a lot of money and man-hours and focus your efforts exclusively on fine-tuning the model, tuning hyperparameters, as well as role-based prompting (the technique of making the right query on the model), which will lead to a significant increase in performance.

After comparing the models, a service was created that had several endpoints and performed the following functions:

- Paraphrasing with enhancement of written text – beautify.
- Translation from English to German.
- Translation from American English to British in one-sentence and multi-sentence mode.
- Checking the grammar of the written text.

Hyperparameters and their configuration. One of the most important things when working with large language models is hyperparameters, which can be adjusted to improve the performance of the pipeline. The most common hyperparameters are the following:

- `input_ids`:
 - Description: A tensor of input token identifiers representing the input text to be generated.
 - Type: `torch.Tensor`.
 - Example: `input_ids`.
- `temperature`:
 - Description: A temperature parameter to control the degree of randomness in the lasing. A high temperature (greater than 1.0) makes the lasing more random, while a low temperature (less than 1.0) makes it more deterministic.
 - Type: `float`.

- Example: temperature=0.7.
- repetition_penalty:
 - Description: Penalty for repeating tokens. A value greater than 1.0 increases the likelihood of rejecting repeated phrases.
 - Type: float.
 - Example: repetition_penalty=1.2.
- num_return_sequences:
 - Description: The number of alternate generation sequences that will be returned.
 - Type: int.
 - Example: num_return_sequences=3.
- no_repeat_ngram_size:
 - Description: The maximum n-gram size to be prevented during generation (no repetitions).
 - Type: int.
 - Example: no_repeat_ngram_size=2.
- num_beams:
 - Description: The number of beams to generate. A higher number of beams can improve the quality of generation, but slow down the process.
 - Type: int.
 - Example: num_beams=5.
- num_beam_groups:
 - Description: Number of beam groups. Ray groups help control the diversity of lasing within each group.
 - Type: int.
 - Example: num_beam_groups=3.
- max_length:
 - Description: The maximum length of the generated sequence (including input tokens). Limits the length of the generation.
 - Type: int.
 - Example: max_length=100.
- diversity_penalty:
 - Description: A parameter that affects the degree of diversity of the sequences generated. A value greater than 1.0 increases variety.
 - Type: float.
 - Example: diversity_penalty=1.5.

Changing each of these parameters can affect the text generation process. Changing the temperature towards higher values makes the generation more random and creative, but can also result in less crisp and less cohesive text. Increasing repetition_penalty can help reduce repetitions in generation. Increasing num_return_sequences gives more alternative generation options. Increasing num_beams and num_beam_groups improves generation quality and variety, but can also slow down the process. Increasing the max_length limits the length of the generation. Changing diversity_penalty affects the degree of diversity of the alternate sequences generated.

Evaluate the performance and quality of models. Let's take a look at how models work on test input data. Figure 2 illustrates how the error correction model works.

Input	Output
See you s00n	See you soon!
I is a trainee at iTechArt Group right now.	I am a trainee at iTechArt Group right now.
I have ability of writing, creating content, to build communication.	I have the ability to write, create content, and build communication.

Figure 2. How the Error Correction Model Works

Figure 3 shows how the American English to British translation model works.

Input	Output
The intermission between graduation and finding the first job was 3 months.	The interval between graduation and finding the first job was 3 months.
I studied math at school.	I studied maths at school.
I have gotten serious about my future career.	I have got serious about my future career.

Figure 3. How the Translation Model Works

Below are the options for the text rewriting model. According to the design, the model should give us three different versions of the rewritten text per sentence. In particular, the model should work well regardless of the length of the input text. This is illustrated in Figures 4 and 5.

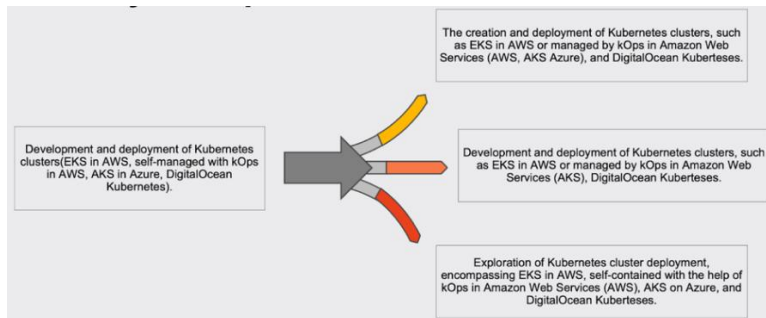


Figure 4. How the text rewriting model works

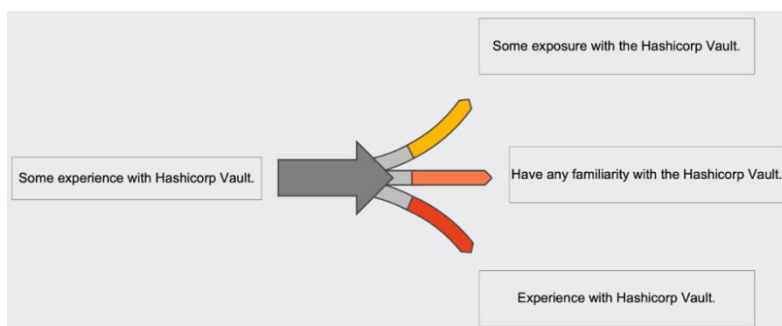


Figure 5. The Single Sentence Text Rewriting Model Works

Disadvantages. However, like everything based on neural networks, the current one has drawbacks. Because it is impossible to absolutely completely learn a model and make it flawless. In our specific examples, the easiest way to see this is in the text rewrite model: since all text-2-text models work on the principle of building text based on maximum probability, there are often mishaps (Fig. 6).

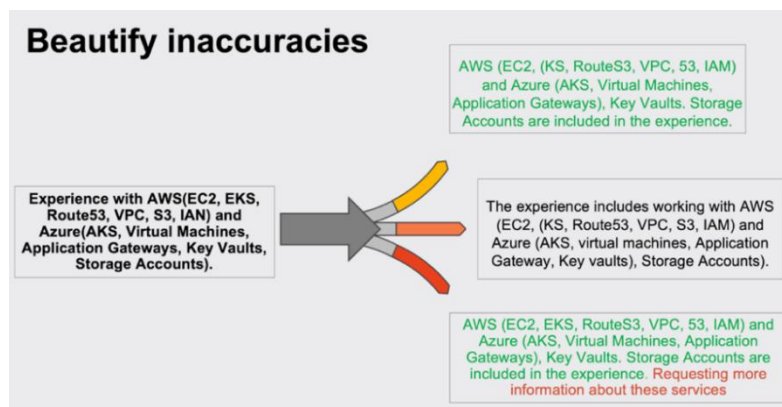


Figure 6. Example of possible errors

The same applies to other models, as, for example, when translating into another language, the names of programming languages can be translated, like python, although in our case this is incorrect. All such cases need to be corrected, which can be done by retraining the model on specific data, for example, by taking the model of a German translator who learned the general base of the language, and then learning it in words from the everyday life of an IT specialist. So we used transfer training.

Transfer learning is a machine learning technique that allows you to reuse a model trained on one task and fine-tune it for a related but different task.

There are different approaches to transfer learning, but one of them, finetuning, is particularly widespread. In this approach, the team takes a pre-trained model and removes or retrains the last layers of that model to focus on a new, similar task.

The advantages of such a model are the use of an existing knowledge base, increased model performance, and low cost.

Performance. Speaking of performance, it's worth noting that we've achieved good performance of the pre-trained models by breaking down the task into smaller ones, namely, in the places of our application, where possible, we send single applications to the service with models, for example, a certain record of the skill of the developer who then passes this model, which allowed us to minimize the response time of the model. Figure 7 shows the average result of our model tests.

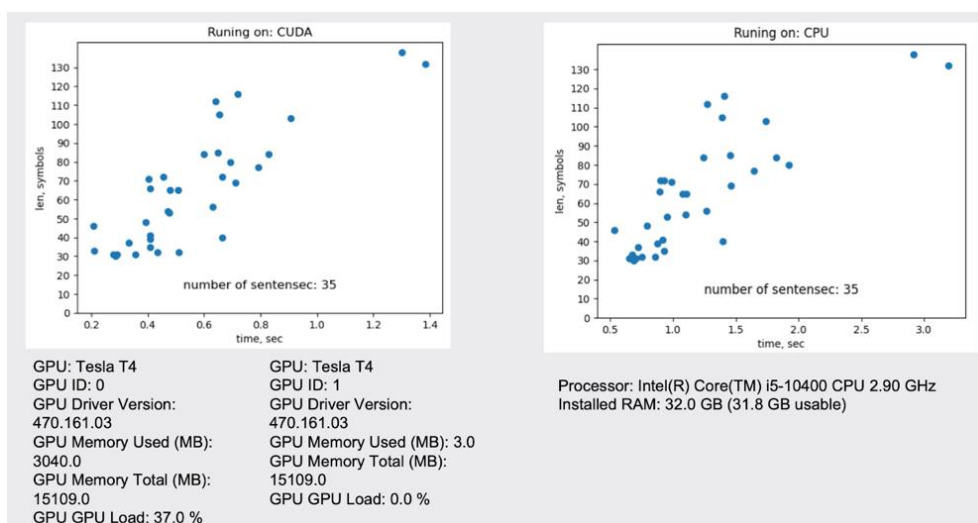


Figure 7. Model Performance

Conclusion. Thus, after adding a service using language models to the existing application, the efficiency of CV writing by the company's employees has increased, the time that an employee spends on a CV has decreased, and most importantly, the quality of the CV has increased and the number of grammatical errors has decreased, and the CV can now be translated into another language at the click of one button, which eliminates the need to sit and translate it by hand.

Reference list

- [1] S. Samarasinghe. Neural Networks for Applied Sciences and Engineering, 2006.
- [2] M.Nielsen. Neural Networks and Deep Learning, 2019.
- [3] C.Olah. Deep Learning, NLP, and Representations, 2014.

author's contribution

Alexey Markov – the head of the study, responsible for setting the tasks and goals of the scientific work.

Maria Zyryanova – study of large language models, implementation and improvement of models responsible for translation from British English into American and grammar checking.

Aleh Asadchy – study of large language models, implementation and improvement of models responsible for translation from English into German and rewriting.

ИССЛЕДОВАНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ ДЛЯ ГЕНЕРАЦИИ ТЕКСТОВ И ИХ ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ

А.Н. Марков

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

М.М. Зырянова

Студентка факультета
компьютерных систем и сетей
Белорусского государственного
университета информатики и
радиоэлектроники.

О.Э. Осадчий

Студент факультета
компьютерных систем и сетей
Белорусского государственного
университета информатики и
радиоэлектроники.

Аннотация. Данное исследование направлено на изучение больших языковых моделей, предназначенных для перевода и генерации текста. В ходе исследования были рассмотрены трансформаторы типа *Text2TextGeneration*, проведен их сравнительный анализ и тестирование различных ситуаций. Большие языковые модели также были внедрены в существующий программный продукт.

Ключевые слова: большие языковые модели, трансформеры, *Text2TextGeneration*, *fine-tuning*, гиперпараметры, нейронные сети, *Big Data*.

UDK: 336.71/004

THEORETICAL ASPECTS OF THE DEVELOPMENT OF BANKS IN THE CONDITIONS OF DIGITAL TRANSFORMATION



D.A. Mirzaev

*Head of the Department of Information Systems and Digital Technologies
Tashkent Institute of Finance,
PhD, Associate Professor
mdilshod@mail.ru*

D.A. Mirzaev

Graduated from Tashkent University of Information Technologies. The area of scientific interests is related to the development of methods and algorithms for constructing information management systems, software reliability, mobile communication systems, organization of educational and research processes in a financial institution.

Annotation. Currently, almost every large bank in Uzbekistan has its own mobile application that helps solve the daily problems of its customers. All mobile banks have basic functions for making basic payments and money transfers. The level of development of the digital bank is increasing year by year, and the range of services offered by the bank is increasing. Now, in the mobile application market, not solving specific problems, but the information that users receive how they use it and control it is becoming more important every day.

Keywords: digital economy, banking system, digital banking, financial service, financial technology, neobank.

Introduction. In the period of development of the digital economy, the banking sector is the flagship of the environment for the technical implementation of digitization. The continuous emergence of technology creates new opportunities for financial customers to interact with each other. Through this opportunity, it provides new information and communication tools for the commercial environment. The growing flow of customers is using more and more communication channels to receive banking services, use new platforms and banking ecosystems. In particular, technology is adapting to changes in people's behavior, resulting in more efficient and cheaper solutions for business development. The beginning of the digital transformation in our country began to be implemented in the banking sector, which is the central component of the economy. The main reasons are, firstly, the banking system is the most sensitive and adaptive system to external factors, which allows to improve the process of providing banking services, develop and introduce banking services by analyzing the existing innovative experience in this field. Digitization of banking activities has become an important step in the development of the modern economy. Credit organizations are introducing digital technologies into their operations, enabling them to be more successful and competitive. Globally, the evolution of banking services has gone through several stages, which contributed to the transition to a digital model of banking services (Fig. 1) [2].

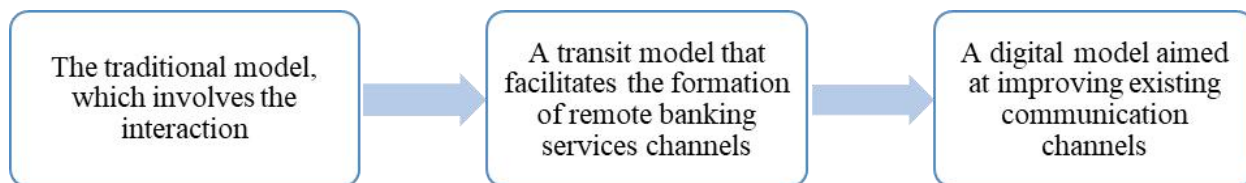


Figure 1. Models of banking services

According to experts, the digital transformation of the banking sector in the Republic of Uzbekistan is characterized by the development of financial technologies that change the system of providing financial services in the aspects of data analysis, automation of commercial processes and distribution of human resources between banks. Since digitization of banking activity is one of the most important processes in the country's economy, the strategy for reforming the banking system of the Republic of Uzbekistan for 2020-2025 was developed [1]. In accordance with the tasks defined in the strategy of further development of the banking system, the following main directions can be singled out:

- regulation of data exchange;
- regulation of ecosystems;
- regulation of payment service providers without the participation of banks;
- improvement of electronic cooperation between market participants, the state, citizens and commerce;
- Development of a number of infrastructure projects;
- Development of digital sum platforms.

The stages of transformation of the banking system are shown in the figure below (Fig. 2).

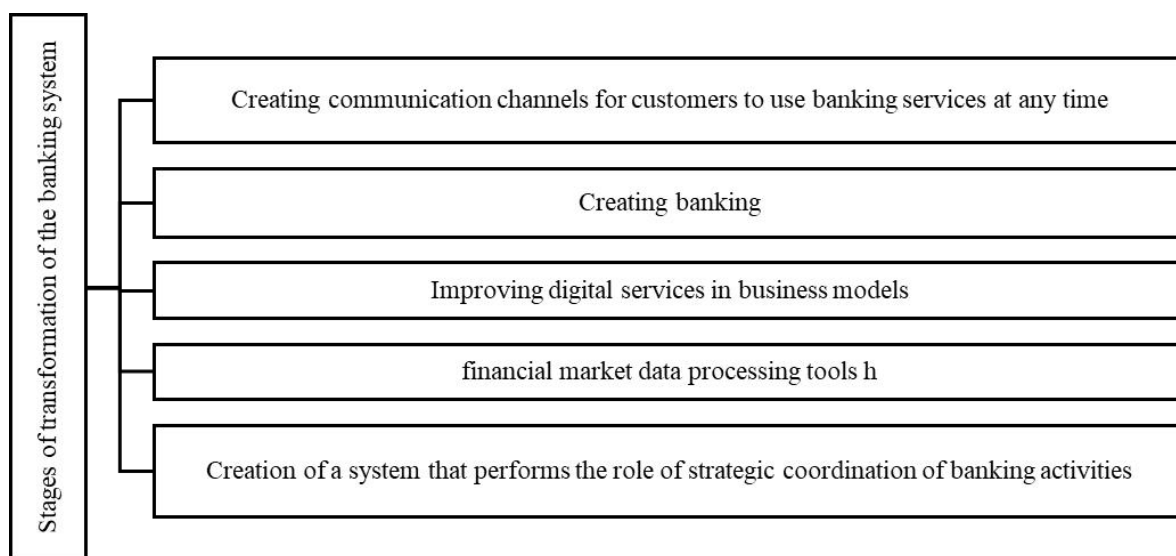


Figure 2. Steps of banking system transformation

It should be said that in the current environment, not every commercial bank could withstand competition or reach a new high level of service and sales of banking products and services. The introduction of digital technologies into the banking sector will directly change the role of banks, which are increasingly focused on personalization and mobility. The modern banking system is characterized by the expansion of the boundaries of the provision of digital services, which directly leads to the formation of new approaches to the management of commercial banking processes.

Globalization and integration processes taking place in the world economy over the last ten years have significantly changed the level of development of the financial market and banking sector. Because of this, they switched to innovative ways of operating. At this stage of development, the banking sector is undergoing changes necessary for its normal functioning in modern reality, and this is due to the influence of the following global factors:

- the evolution of consumer habits and customer preferences;
- development of financial technologies and innovations and their rapid introduction;
- Increased competition due to financial technologies.

In the banking system, there are almost no operations and services that are not implemented by financial technologies. In addition, such services have a number of advantages in the form of low cost, speed of delivery and convenience. That is why there is a tendency in Uzbekistan and in the world that banks with traditional banking services are losing their customers, because new financial technologies, for example, "neobanks", are creating more convenience for consumers in making transactions. The main components of the digital model of banking services for customers of Uzbekistan banks can be seen in the following figure (Fig. 3).

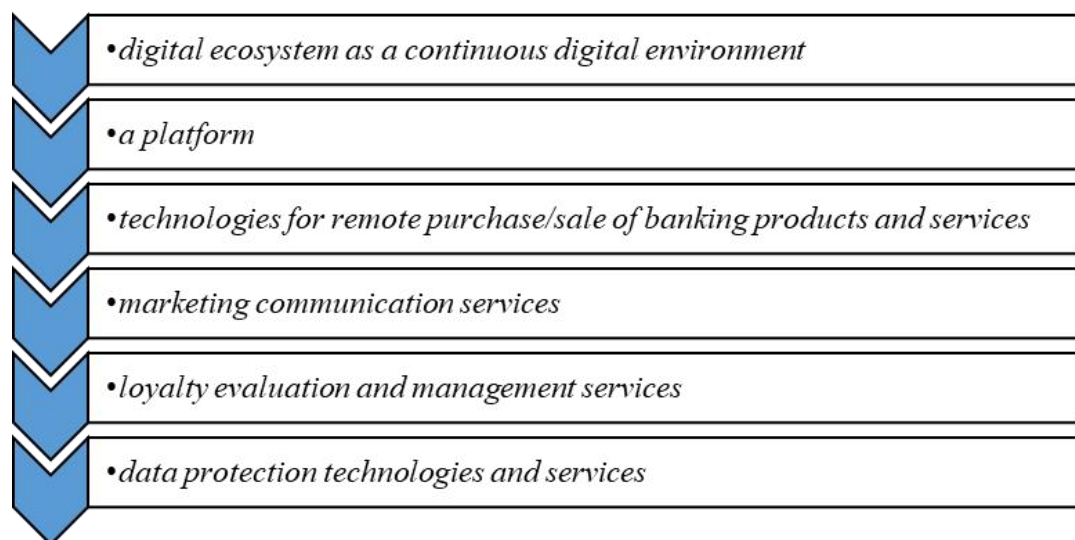


Figure 3. Composition of the digital model of banking services for corporate clients by banks

Based on the above, it seems appropriate to propose the idea of creating a neobank model focused on serving corporate clients. Creation of the Neobank model offers a new approach in the field of creation of digital communication channels with corporate clients to provide technological banking products online based on the use of machine learning technology, artificial intelligence systems, big data technology, predictive analytics, Internet of Things, Internet technologies is enough. Digital banks are also classified as social in part because with the transition to the digital space, banks continue to connect with customers using social networks. The following positive aspects of the bank's transition to the digital space are also highlighted [3]:

- the number of cash transactions is decreasing, more transactions are switching to cashless form, which makes them more transparent;
- branches reduce the burden of sales and transaction operations, and they can focus more on advising on more complex products;
- The intangible nature of the products offered by banks makes it easy to organize their sale via the Internet.

Newly established banks or small banks with an underdeveloped branch network can move to the digital space much faster, whereas for large banks, the transition to digital banking is more laborious, time-consuming and expensive. Therefore, it is more efficient for some large banks to

create a new digital bank, which allows them to retain their existing customers and attract new customers who want to use a more convenient online service. However, in addition to many advantages, digitization of the banking sector has significant disadvantages. One of them is that the further development of digitalization of the banking sector requires large investments, and such funds are not always available in small financial institutions. Larger banks, unlike smaller banks, have more capital to hire the best employees, buy new technology, or introduce artificial intelligence and biometric technologies into their organizations. In this regard, it is easier for large banks to compete with smaller players in the field of digitization of the banking sector. Therefore, banks with large capital can exclude from the market, especially small financial institutions that have just started to develop [4].

Another disadvantage is that people of retirement age are not yet ready to give up personal communication with employees of the banking sector. The reason is that people of retirement age prefer not to use digital products unless necessary, and prefer to visit offices live rather than interact with robots. In addition, many lack the financial literacy skills necessary to access digital services. Unfortunately, not everyone understands the importance of digital transformation. For many, the introduction of new technologies is a threat, as the risk of data leakage and cybercrime increases every year. Digitization of banks is being implemented in two directions, that is, by introducing product and process innovations [5].

The first path of development is based on the release of innovative banking products, which include a credit broker service that allows you to fill out an online application for a consumer loan or mortgage without the involvement of specialists, biometric identification that allows confirmation, remote banking operations, cashback services, permanent implements bonus programs and other types of services for customers.

The second direction of digitization is process innovation, which is primarily aimed at minimizing costs related to the support of banking operations. The use of artificial intelligence involves the optimization of banking services by reducing the intermediary relationship with the customer and providing a wide range of services to meet the needs of all segments of the population. [6].

Summary. At the same time, the digital transformation of the economy is considered one of the priority tasks of the country's development strategy. Additionally, the impact of the 2020 global coronavirus disease pandemic (COVID-19) and the resulting acceleration of digital technologies has forced the banking ecosystem to restructure around digital experiences like never before. The definition of necessary and sufficient features in digital banking products is changing rapidly, and the competition for profitable primary relationships has intensified. It is possible to highlight the main advantages of using modern Internet technologies in the banking sector:

- save time;
- the ability to carry out transactions remotely without going to the bank in person;
- Immediate response to changes in the financial market, because bank customers can perform the necessary operations with their accounts at any time of the day.

References

- [1] O‘zbekiston Respublikasi Prezidentining 2020-yil 12-maydagi «2020 – 2025-yillarga mo‘ljallangan O‘zbekiston Respublikasining bank tizimini isloh qilish strategiyasi to‘g‘risida» PF-5992-sonli Farmoni.
- [2] Панова Г.С. Эволюция традиционных банков в новых реалиях финансовых экосистем // Банковские услуги. – 2021. – № 5. – С. 2-8.
- [3] Юсупбеков Н.Р., Гулямов Ш.М., Усманова Н.Б., Мирзаев Д.А. Тенденции развития систем управления технологическими процессами и производствами: обзор в контексте исследований и разработок // Промышленные АСУ и контроллеры. - 2017. - № 9. - С. 3-9.
- [4] Greenspan A. The payments system in transition // Federal Reserve Payments System Development Committee 2003 Conference. BIS Review 46/2003. Washington, DC, 2003. - P. 4.

[5] Ихамова Y.S., Mirzayev D.A., Mo‘minov B.B. «Tijorat banklarida raqamli texnologiyalardan foydalanish istiqbollari»: Monografiya. - Toshkent, 2023. – 117 p.

[6] Бердиназаров З.У., Исаев А.Д. Цифровая валюта и её внедрение в Узбекистане. Сборник международной научно-практической конференции. Актуальные вопросы современной науки и образования – 2022. – С. 168.

ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ РАЗВИТИЯ БАНКОВ В УСЛОВИЯХ ЦИФРОВОЙ ТРАНСФОРМАЦИИ

Д.А. Мирзаев

*Заведующий кафедрой информационных систем
и цифровых технологий*

*Ташкентского финансового института,
к.т.н., доцент*

Аннотация. В статье освещается текущее состояние цифровой трансформации бизнеса в Узбекистане на примере развития цифровизации банковского сектора. В настоящее время практически каждый крупный банк Узбекистана имеет собственное мобильное приложение, помогающее решать повседневные проблемы своих клиентов. Все мобильные банки имеют базовые функции для совершения основных платежей и денежных переводов. Уровень развития цифрового банка растет с каждым годом, увеличивается спектр предлагаемых банком услуг. На рынке мобильных приложений с каждым днём всё актуальнее становится информация, которую могут получить пользователи, а решение конкретных задач становится вторичным.

Ключевые слова: цифровая экономика, банковская система, цифровой банкинг, финансовый сервис, финансовые технологии, необанк.

УДК 621.865.8:629.055

MOTION CONTROL OF MOBILE ROBOT WITH DIFFERENTIAL DRIVE AND TWO CASTOR WHEELS



Yang Shixin

PhD student of the Department of «Electric drive and automation of industrial installations and technological complexes» of BNTU
yang_shi_xin@163.com



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
anna.velchenko@gmail.com



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
sap@bntu.by
s.pauliukavets@gmail.com

S.A. Pauliukavets

PhD in Technical Sciences, Associate Professor, Head of the of the «Electric Drive and Automation of Industrial Installations and Technological Complexes» Department of the Belarusian National Technical University. The area of scientific interests is related to the development of methods and algorithms for controlling mobile robots.

A.A. Velchenko

PhD in Technical Sciences, Associate Professor of the «Electric Drive and Automation of Industrial Installations and Technological Complexes» Department of the Belarusian National Technical University. The area of scientific interests is related to the development and research of neural network controllers for electric drives.

Yang Shixin

PhD student at the Belarusian National Technical University. The area of scientific interests is related to the study of control algorithms for wheeled mobile robots.

Abstract. This article presents a mathematical model of a motion control system for a nonholonomic mobile robot, an independent suspension with two independent drive wheels and two castor wheels. A model that takes into account the position of a wheeled robot relative to a given path, using transfer design function devices, which allows the use of methods of both classical and modern control theories for the synthesis of control devices.

Keywords: mobile robot, mathematical model, castor wheel, kinematics, control system, motion trajectory.

Introduction. The differential mobile robot is simple in structure, low in cost, relatively easy to analyze and establish the kinematics model, and currently has many applications in the market [1], so this paper chooses the differential mobile robot as the research object. The system input of the differential robot is only the speed of the left and right driving wheels, and the forward, backward, turning and other actions are realized through the speed difference. Usually, in order to improve the balance of the robot, a universal wheel is installed in the front, which only plays a supporting role and does not provide power [2]. The establishment of the robot motion model is a process from bottom to top. The motion of the robot is realized through the action of the wheels, but the wheels also constrain the motion of the robot. Multiple wheels are

connected together through a mechanical structure to form a robot chassis. Therefore, the constraints of each wheel are combined to generate motion constraints on the robot chassis. At this time [3], it is necessary to analyze the kinematics model of the robot.

To improve the design of robots and reduce costs, it is necessary to research and use new, stronger and lighter materials, develop fundamentally new designs, and optimize robot control algorithms. Optimizing a robot control algorithm or developing a new algorithm involves studying the functioning of the robot using a mathematical model. When developing mathematical models, special attention should be paid to modeling the external conditions in which the robot is located, i.e. modeling the load on its drive motors. Many mathematical models have been developed that describe the electrical and mechanical parts of electric drives of mobile robots and other mechanisms [4–6]. This paper sets the task of controlling the movement of a mobile robot based on its mathematical model. When developing a mathematical model, it is necessary to take into account the external conditions in which the robot is located and moves.

Difference robot coordinate transformation. There are multiple coordinate systems in the robot system, including the world coordinate system ($X_W O_W Y_W$), robot coordinate system ($X_R O_R Y_R$), Sensor coordinate system ($X_S O_S Y_S$). As shown in Figure 1. The environment in which the robot moves is called the world coordinate system, which is fixed and used to describe the global information of the robot. The origin of the robot coordinate system is the midpoint of the two driving wheels, and the instantaneous orientation of the robot is the positive direction of the coordinate axis. When the robot calculates the turning radius, it also calculates the distance from the instantaneous center of rotation to the midpoint of the driving wheels. The robot coordinate system is used to describe the robot's own information, and it is also the conversion medium from the sensor coordinate system to the world coordinate system. The information collected by the sensor is based on the sensor coordinate system. For example, the data information detected by the lidar is based on the lidar coordinate system.

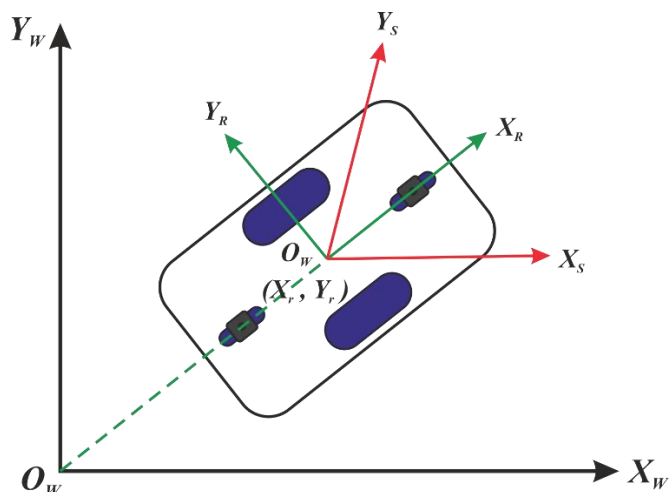


Figure 1. Schematic diagram of chassis movement of differential mobile robot

The origin position and orientation of the robot coordinate system and the sensor coordinate system change with the movement of the robot, so the robot posture and sensor data information cannot be directly applied, but should be changed to the world coordinate system. The transformation from the robot coordinate system to the world coordinate system is related to the pose of the robot at the current moment. The sensor is fixed on the robot through a mechanical structure, so the transformation relationship from the sensor coordinate system to the robot coordinate system is determined by the installation position of the sensor.

The origin of the robot coordinate system $X_R O_R Y_R$ is O_R . The pose in the world coordinate system $X_W O_W Y_W$ is $P_{WR}(x_r, y_r, \theta_r)$. The conversion matrix from the robot coordinate system $X_R O_R Y_R$ to the world coordinate system $X_W O_W Y_W$ is:

$$T_{WR} = \begin{bmatrix} \cos(\theta_r) & -\sin(\theta_r) & x_r \\ \sin(\theta_r) & \cos(\theta_r) & y_r \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (1)$$

Of which:

$$R = \begin{bmatrix} \cos(\theta_r) & -\sin(\theta_r) \\ \sin(\theta_r) & \cos(\theta_r) \end{bmatrix} \quad (2)$$

$$t = \begin{bmatrix} x_r \\ y_r \end{bmatrix} \quad (3)$$

The pose of the origin of the sensor coordinate system in the robot coordinate system is $P_{RS} = (x_{rs}, y_{rs}, \theta_{rs})$. The coordinates in the world coordinate system are as follows:

$$P_{WS} = P_{RS} T_{WR} = (x_{rs}, y_{rs}, \theta_{rs}) \cdot \begin{bmatrix} \cos(\theta_r) & -\sin(\theta_r) & x_r \\ \sin(\theta_r) & \cos(\theta_r) & y_r \\ 0 & 0 & 1 \end{bmatrix} = (x_{ws}, y_{ws}, \theta_{ws}) \quad (4)$$

From this, the conversion matrix of the sensor coordinate system $X_S O_S Y_S$ to the world coordinate system $X_W O_W Y_W$ can be obtained:

$$T_{WS} = \begin{bmatrix} \cos(\theta_{ws}) & -\sin(\theta_{ws}) & x_{ws} \\ \sin(\theta_{ws}) & \cos(\theta_{ws}) & y_{ws} \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

Robot kinematics model. The pose of the robot in the environment is described as (x, y, θ) . There are three degrees of freedom. The input of the system is only the speed of the left and right wheels (v_L, v_R). So the differential mobile robot is an underactuated system, and its motion is coupled. Figure 2 is a schematic diagram of the chassis motion of the differential mobile robot studied in this paper.

What is known in the above figure is the movement speed of the left and right wheels v_L, v_R . It can be calculated by the photoelectric encoder. The distance between the centers of the two driving wheels is $2d$, which has been determined during installation and can be obtained by measurement. Other information of the robot can be deduced from the known information through the establishment of the kinematics model. Assume that the linear and angular velocities of the robot chassis are v, ω . The angular velocity of the left and right wheels is ω_L, ω_R . The instantaneous turning radius of the robot when moving is r . The movement of the differential mobile robot at any moment can be regarded as a circular arc movement. The entire robot system is a rigid body connected uniformly. When the robot moves, the angular velocity of the left and right wheels is equal to the overall angular velocity of the robot.

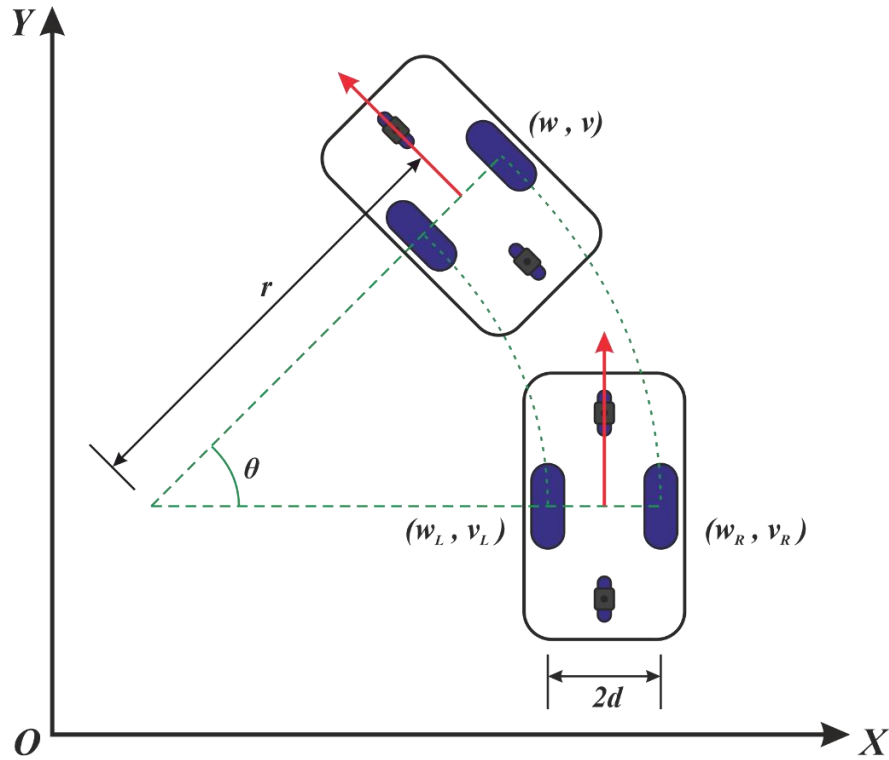


Figure 2. Schematic diagram of differential chassis movement

$$\omega = \omega_L = \omega_R \quad (6)$$

From the relationship between angular velocity and linear velocity, we can get:

$$\frac{v_L}{r-d} = \frac{v_R}{r+d}$$

$$v_L(r+d) = v_R(r-d) \quad (7)$$

$$r(v_R - v_L) = (v_R + v_L)d$$

So the formula for the instantaneous radius of rotation is as follows:

$$r = \frac{(v_R + v_L)d}{v_R - v_L} \quad (8)$$

Calculated from the angular velocity formula:

$$\omega = \frac{v_R}{r+d} \quad (9)$$

$$r+d = \frac{(v_R + v_L)d}{v_R - v_L} + d = 2 \frac{v_R d}{v_R - v_L} \quad (10)$$

Substitute formula (10) into formula (9) to get the angular velocity formula:

$$\omega = \frac{v_R - v_L}{2d} \quad (11)$$

Linear speed of differential robot chassis:

$$v = \omega * r = \frac{v_R - v_L}{2d} \frac{(v_R + v_L)d}{v_R - v_L} = \frac{v_R + v_L}{2} \quad (12)$$

The kinematics model is expressed in matrix form as follows:

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2d} & \frac{1}{2d} \end{bmatrix} \begin{bmatrix} v_L \\ v_R \end{bmatrix} \quad (13)$$

The differential drive robot determines three different motion states according to the speed difference between the left and right wheels, as shown in Figure 3.

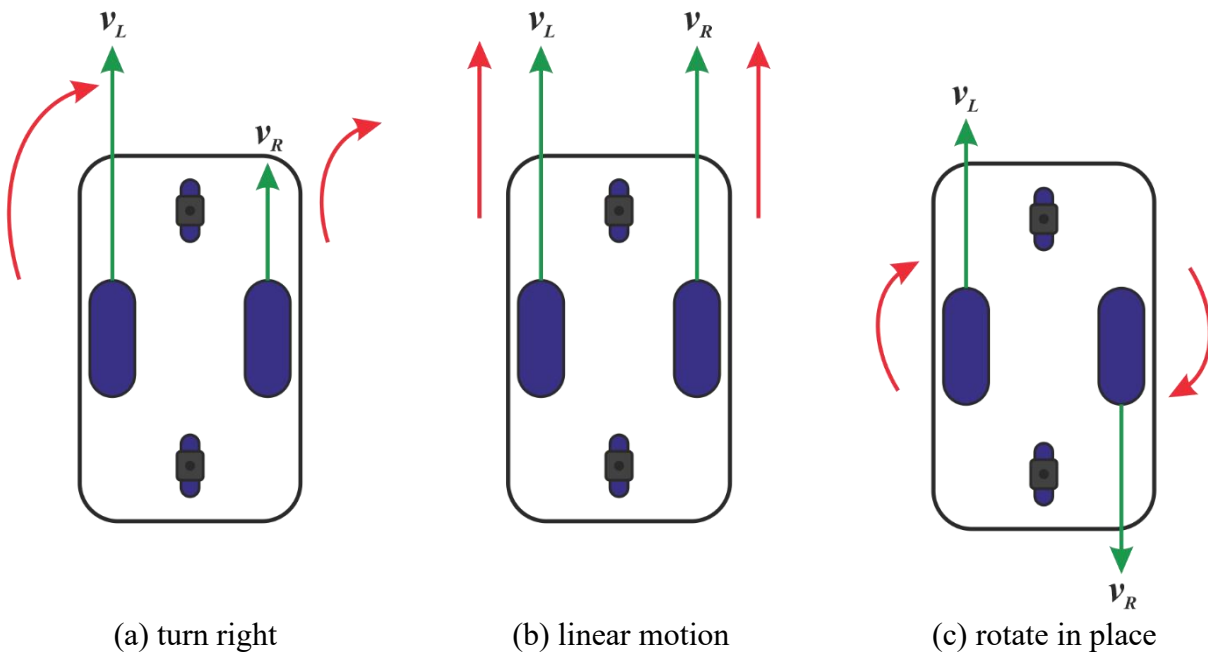


Figure 3. Three motion states of the differential robot

When $v_L > v_R$, make a circular motion and turn right.

When $v_L < v_R$, true life.

When $v_L = v_R$, make a straight line motion.

When $v_L = -v_R$, rotate in situ with the center point of the left and right wheels.

Conclusion. The control of a nonholonomic mobile robot with a differential drive and two castor wheels along a reference curve is presented, where the controller consists of two parts: feedforward control and feedback control. The former uses robot inverse kinematics to calculate feedforward inputs based on a reference curve, while the latter removes the effects of noise, disturbances, and initial state errors. The presented example proves that a mobile robot can follow a desired reference path according to a given speed profile with satisfactory accuracy. Thus, the presented model is useful for a preliminary study of the designed motion control laws of a mobile robot.

References

- [1] Berman S., Schechtman E., Edan Y. Evaluation of Automatic Guided Vehicle Systems // Robotics and Computer-Integrated Manufacturing, 2008. doi: 10.1016/j.rcim.2008.02.009.
- [2] Siegwart R., Nourbakhsh I. R., Scaramuzza D. Introduction to autonomous mobile robots. Publishing house of Massachusetts Institute of Technology, 2004. 321 p.

[3] Xiang Yuyun. Research on indoor robot navigation control method based on SLAM [D]. Northwest A&F University, 2021.

[4] R. X. Xu. Path planning of mobile robot based on multi-sensor information fusion[J]. EURASIP Journal on Wireless Communications and Networking, 2019, 2019(1): 1-8.

[5] Zeng Wentе, Su Jianbo. A network robot system based on distributed intelligence [J]. Robotics, 2009.

[6] Tian Ye, Chen Hongwei, Wang Fasheng, et al. A Review of SLAM Algorithms for Indoor Mobile Robots J. Computer Science, 2021, 48(09): 223-234.

Авторский вклад

Все авторы внесли равноценный вклад в написании статьи.

УПРАВЛЕНИЕ ДВИЖЕНИЕМ МОБИЛЬНОГО РОБОТА С ДИФФЕРЕНЦИАЛЬНЫМ ПРИВОДОМ И ДВУМЯ ПОВОРОТНЫМИ КОЛЕСАМИ

Ян Шисинь

*Аспирант кафедры
«Электропривод и
автоматизация промышленных
установок и технологических
комплексов» БНТУ*

А.А. Вельченко

*Доцент кафедры
«Электропривод и
автоматизация промышленных
установок и технологических
комплексов» БНТУ, кандидат
технических наук, доцент*

С.А. Павлюковец

*Заведующий кафедры
«Электропривод и
автоматизация промышленных
установок и технологических
комплексов» БНТУ, кандидат
технических наук, доцент*

Аннотация. В данной статье представлена математическая модель системы управления движением негोलонномным мобильным роботом, имеющего шасси с двумя независимыми ведущими колесами и двумя кастор колесами. Модель, учитывающая положение колесного мобильного робота относительно заданного пути, построена с использованием аппарата передаточных функций, что позволяет применять методы как классической, так и современной теории управления для синтеза устройств управления.

Ключевые слова: мобильный робот, математическая модель, поворотное колесо, кинематика, система управления, траектория движения.

УДК 004.021:004.75

ИННОВАЦИОННЫЕ ТЕХНОЛОГИИ ОБУЧЕНИЯ ДЛЯ ТРЕНИРОВКИ МЕТКОСТИ СТРЕЛБЫ НА ПРИМЕРЕ VR-СИМУЛЯТОРА «КОНЦЕНТРАЦИЯ МАХ»



Е.А. Алетунович
Учащийся Учреждения
образования «Национальный
детский технопарк»,
учащийся
ГУО «СШ №16 имени
П.М.Машерова г.Лиды»



О.Д. Лоская
Учащийся Учреждения
образования «Национальный
детский технопарк»,
учащаяся ГУО «СШ №46 г.
Витебска им. И. Х.
Баграмяна»



М.С. Ильясова
ассистент кафедры
инженерной психологии и
эргономики БГУИР, магистр



Усенко Ф.В.
Инженер–программист
кафедры инженерной
психологии и эргономики
БГУИР, магистр



Л.Р. Коркин
Заведующий лабораторией
виртуальной и дополненной
реальности Национального
детского технопарка,
магистр технических наук



А.М. Прудник
доцент кафедры инженерной
психологии и эргономики
БГУИР, кандидат
технических наук, доцент
aleksander.prudnik@bsuir.by

Е.А. Алетунович

Обучается в школе г. Лиды. Область научных интересов связана с технологиями виртуальной и дополненной реальности.

О.Д. Лоская

Обучается в школе г. Витебска. 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.

Усенко Ф.В.

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой манипуляторов для дистанционного управления.

Коркин Л.Р.

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. Статья исследует потенциал виртуальной реальности (VR) в области обучения стрельбе, особенно в контексте допризывной подготовки и тренировки военно-патриотических ценностей у молодежи. Рассматривается разработка сценариев игровых симуляторов тира в VR, представляющих собой эффективные инструменты для повышения меткости, реакции и стрелковых навыков. Обсуждаются различные типы стрельбищ, включая стационарные и движущиеся мишени, а также скрипты управления игровым процессом. Исследование акцентирует внимание на потенциале виртуальной реальности в улучшении качества допризывной подготовки и повышении эффективности обучения в области стрельбы и оружия.

Ключевые слова: VR, Unity, виртуальная реальность, обучение стрельбе, симуляторы тира, игровые механики, допризывная подготовка, образовательные технологии, виртуальные симуляторы.

Введение. В современном обществе одной из важнейших задач является эффективное обучение и подготовка молодежи, особенно в контексте допризывной подготовки. Отсутствие современных средств и методов обучения может ограничивать возможности как в обучении навыкам стрельбы, так и в формировании военно-патриотических ценностей у молодежи [1].

Игровые технологии, особенно в контексте виртуальной реальности (VR), представляют собой мощный инструмент, обладающий значительными воспитательными возможностями. В настоящее время игры в виртуальной реальности обладают огромным потенциалом и активно внедряются в сферу образования [2].

Актуальность данного исследования обусловлена необходимостью улучшения качества допризывной подготовки, в частности, повышения меткости стрельбы и организации обучения с использованием современных технологий, таких как виртуальная реальность.

Основная часть. Виртуальные симуляторы тира нашли широкое применение в различных областях. Они используются для обучения военных и правоохранительных органов, тренировки меткости и стрелковых навыков, оценки кандидатов на должности, повышения безопасности при обращении с оружием, а также для научных исследований в области оружия и стрельбы. Все это позволяет улучшить качество обучения и безопасность в сфере стрельбы и оружия.

Разработка сценария игры. Игра виртуальной реальности представляет собой тренажёр стрельбы, который ориентирован как на тренировку, так и на проведение соревнований в условиях виртуального пространства.

Первое стрельбище создано в виде ряда мест для стрельбы с интерактивными мишенями (рис.1). Расстояние до мишеней может быть изменено в процессе игры, что позволяет адаптировать сложность тренировки под разные уровни игроков. Мишени двигаются по предварительно заданной траектории, на которой располагаются различные препятствия, такие как стены или барьеры, что усложняет задачу попадания. Каждое место для стрельбы оборудовано монитором, который показывает количество заработанных очков и точное расположение попаданий, что обеспечивает обратную связь и возможность анализа результатов. Подобное стрельбище призвано развивать не только меткость, но и реакцию и скорость принятия решений в условиях изменяющейся среды.

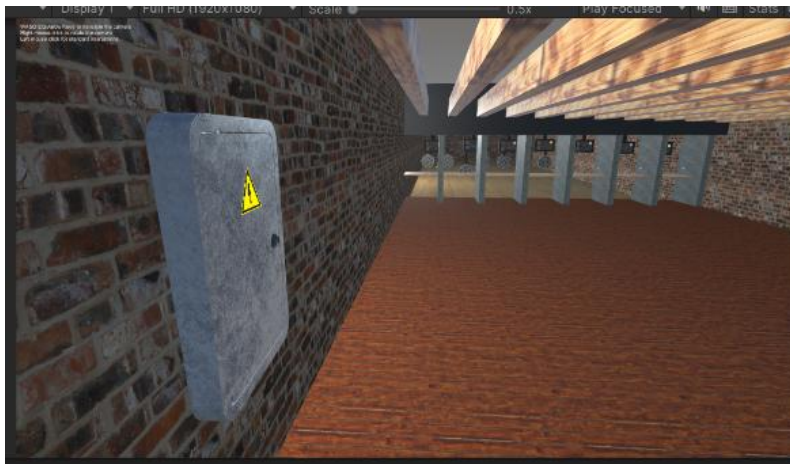


Рисунок 1. Первое стрельбище

Второе стрельбище представляет собой набор движущихся мишеней, двигающихся по заранее заданной траектории, как показано на рисунке 2. Это создает дополнительную сложность, поскольку игроку необходимо адаптироваться к движущейся мишени и стрелять в ней в нужный момент, чтобы достичь максимального результата. Подобные тренировки способствуют развитию ловкости, координации и реакции игрока.

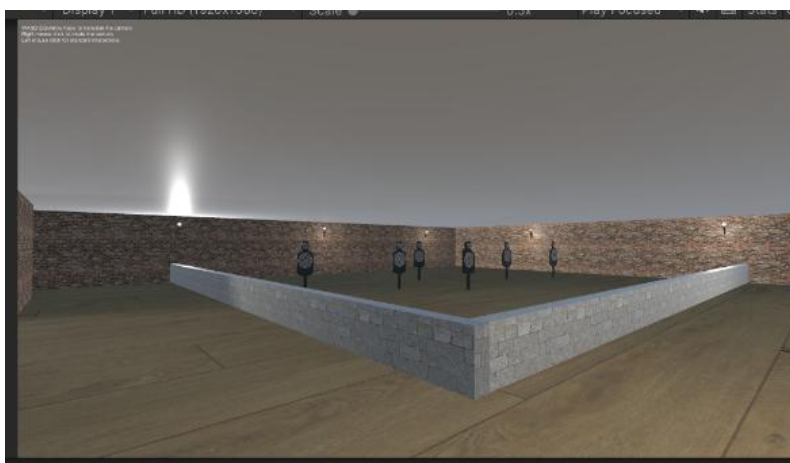


Рисунок 2. Второе стрельбище

Третье стрельбище предназначено для отработки стрельбы по статичным мишеням, размещенным на различной высоте. Здесь дальность стрельбы зависит от расстояния, на котором находится сам игрок. Такой подход к тренировке позволяет игроку самостоятельно регулировать уровень сложности и адаптировать тренировку под свои потребности.

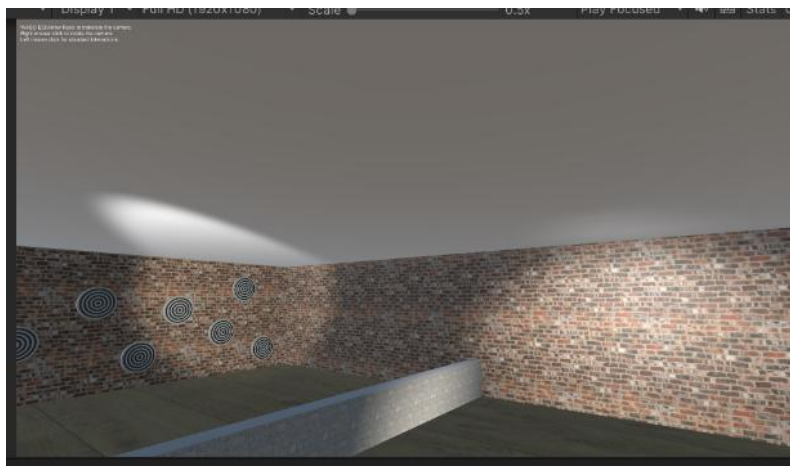


Рисунок 3. Третье стрельбище

После разработки концепции игры и дизайна, включающей поиск и анализ референсов тира как в игровых приложениях, так и в реальной жизни, были созданы трехмерные модели [3]. Последующая текстуризация моделей и их интеграция в программное обеспечение Unity выполнялись с учетом полученных результатов. Несмотря на высокую степень сложности и трудоемкости данного этапа работы, существенным преимуществом в использовании собственных трехмерных моделей является упрощение процесса подбора цветовой гаммы и определения размеров объектов.

Некоторые из объектов такие, как мишени, стены, управляющие элементы, представлены на рисунках 4-5.

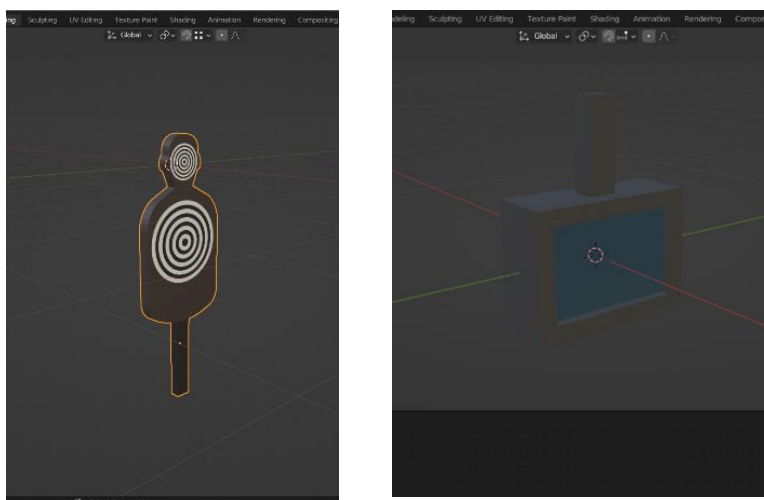


Рисунок 4. 3D-модели объектов взаимодействия

Они выполняют такие функции, как обеспечение удобства игрового процесса, определение пространственных границ игрового мира, добавления динамики и ощущения вызова игровому процессу.

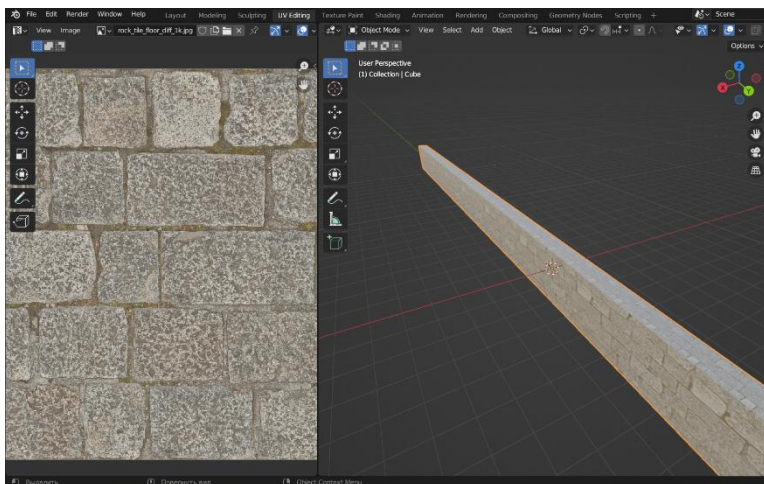


Рисунок 5. 3D-модели стены для ограничения зон

Разработка скриптов игры. Для организации игрового процесса необходимо написать программный код в виде скриптов [4]. Так, для управления основными аспектами стрельбы и взаимодействия с патронами был разработан скрипт компонента оружия. Этот скрипт представляет собой компонент оружия в игре, который управляет основными аспектами стрельбы и взаимодействия с патронами. В методе «Update» определена логика стрельбы: в зависимости от выбранного в настройках режима ввода, при нажатии кнопки выстрела вызывается метод «StartShoot», который начинает стрельбу (рис. 6). Этот метод запускает стрельбу в одиночном или автоматическом (с задержкой между выстрелами) режиме в зависимости от настроек оружия.

```
private void Update()
{
    if (GameSettings.Settings.isControllerInput)
    {
        if (GameSettings.Settings.fireAction.GetStateDown(SteamVR_Input_Sources.Any))
            StartShoot();
    }
    else{
        if (Input.GetKeyDown(GameSettings.Settings.fireButton))
            StartShoot();}

    void StartShoot()
    {
        if (weaponSettings.shootingType == ShootingType.Single)
            Shoot();
        else
            StartCoroutine(StartShooting());
    }
}
```

Рисунок 6. Скрипт для осуществления стрельбы

Скрипт, представленный на рисунке 7, реализует поведение движущейся мишени в контексте виртуального тира, предназначенного для тренировки меткости стрельбы. Представляет собой компонент виртуальной среды, обеспечивающий анимацию движения и взаимодействие при попадании в мишень. Методы *MoveToEnd* и *MoveToStart* отвечают за анимированное перемещение мишени между начальной и конечной позициями, используют библиотеку *LeanTween* для создания плавного перемещения. Метод *GetScore* вызывается при попадании в мишень и определяет, была ли мишень поражена ранее.

Также он инициирует анимацию вращения мишени при попадании, определяет, в какую часть мишени было произведено попадание, и возвращает соответствующее количество очков.

```
public override int GetScore(Vector3 hitPoint, Transform centerPoint =
null, List<int> scoresList = null)
{
    if (!_isKnocked)
        return 0;
    _isKnocked = true;
    LeanTween.value(knockedAnimation.targetGameObject.gameObject,
knockedAnimation.startValue, knockedAnimation.targetValue,
knockedAnimation.duration).setEase(knockedAnimation.curve)
        .setOnUpdate((float value) =>
        {
            knockedAnimation.targetGameObject.rotation =
Quaternion.Euler(knockedAnimation.targetGameObject.rotation.x,
knockedAnimation.targetGameObject.rotation.y, value);
        });
    float distanceHead = Vector3.Distance(hitPoint, centerHead.position);
    float distanceBody = Vector3.Distance(hitPoint, Center.position);
    int score = 0;
    if (distanceHead < distanceBody)
        score = base.GetScore(hitPoint, centerHead, headScores);
    else
        score = base.GetScore(hitPoint, Center);
    FindObjectOfType<ToysShop>().AddBalance(score);
    return score;
}
private void MoveToEnd()
{
    LeanTween.value(gameObject, startPosition.position,
endPosition.position, moveingDuration).setOnUpdate(
    (Vector3 value) =>
    {
        Vector3 newPosition = new Vector3(value.x,
transform.position.y, value.z);
        transform.position = newPosition;
    }).setOnComplete(MoveToStart);
}
private void MoveToStart()
{
    LeanTween.value(gameObject, endPosition.position,
startPosition.position, moveingDuration).setOnUpdate(
    (Vector3 value) =>
    {
        Vector3 newPosition = new Vector3(value.x,
transform.position.y, value.z);
        transform.position = newPosition;
    }).setOnComplete(MoveToEnd);
}
```

Рисунок 7. Скрипт реализации движущейся мишени

Управление движением игрока с учетом возможности отталкивания от стен реализуется с помощью соответствующего скрипта (рис.8). В основе его функциональности лежит алгоритм линейной интерполяции, обеспечивающий плавное и постепенное перемещение игрового объекта, представляющего игрока, от его текущей

позиции к новой целевой позиции. Параметр `speed` определяет скорость движения, а `minDistanceToStop` устанавливает минимальное расстояние, при котором движение прекращается, для точной остановки игрока у стены.

```
using UnityEngine;
public class PlayerMovement : MonoBehaviour
{
    [SerializeField] private float speed;
    [SerializeField, Range(0.05f, 1f)] private float minDistanceToStop;
    private Vector3 _newPosition;
    private bool _moveing;
    public void MovePlayer(Vector3 newPosition)
    {
        _newPosition = newPosition;
        _moveing = true;
    }
    private void Update()
    {
        if (!_moveing)
            return;
        transform.position = Vector3.Lerp(transform.position,
        _newPosition, speed);
        if (Vector3.Distance(transform.position, _newPosition) <=
minDistanceToStop)
            _moveing = false;
    }
}
```

Рисунок 7. Скрипт перемещения игрока

Таким образом, данный скрипт обеспечивает надежное и точное управление движением игрока в виртуальном пространстве, что важно для обеспечения комфортного взаимодействия пользователя с игровым миром и предотвращения коллизий с препятствиями.

Заключение. Использование игровых технологий, особенно в контексте виртуальной реальности, демонстрирует значительный потенциал в образовании и воспитании, способствуя улучшению навыков стрельбы, развитию реакции и координации, а также формированию военно-патриотических ценностей. Реализация разнообразных стрельбищ с использованием интерактивных мишеней и движущихся объектов, а также разработка соответствующих скриптов обеспечивают эффективное обучение, при этом создавая комфортное и реалистичное виртуальное окружение для пользователей. Такие инновационные подходы позволяют современным образовательным учреждениям и военным организациям добиться лучших результатов в обучении и подготовке кадров, а также повысить уровень безопасности при обращении с оружием.

Список литературы

[1] Игровые технологии в обучении и воспитании [Электронный ресурс]. Режим доступа: <https://nsportal.ru/shkola/raznoe/library/2022/04/07/statya-igrovye-tehnologii-v-obuchenii-i-vospitanii>. Дата доступа: 02.02.2024.

[2] How to use Virtual Reality to Improve Your Life: Part 1 [Электронный ресурс]. Режим доступа: <https://medium.com/@iamVictorW/how-to-use-virtual-reality-to-improve-your-life-part-1-a0b724fa427c>. Дата доступа: 10.01.2024.

[3] Какие функции редактора blender [Электронный ресурс]. Режим доступа: <https://typerus.ru/kakiye-funksii-redaktora-blender/?ysclid=ljrbcu52o7864163245>. Дата доступа: 11.01.2024

[4] 16. Visual Studio | Microsoft. Режим доступа: <https://visualstudio.microsoft.com> - Дата доступа: 11.02.2024.

Авторский вклад

Авторы внести равноценный вклад.

INNOVATIVE TECHNOLOGIES FOR TRAINING MARKSMANSHIP USING THE EXAMPLE OF THE VR SIMULATOR «CONCENTRATION MAX»

E.A. Aleturovich

Student of the educational Institution «National Children's Technopark», student SEI «Secondary School No. 16 named after P.M.Masherov, Lida»

O.D. Losskaya

Student of the educational Institution «National Children's Technopark», student SEI «Secondary school No. 46 of Vitebsk named after I. H. Baghramyan»

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

Head of the Virtual and Augmented Reality Laboratory of the National Children's Technopark, Master of Technical Sciences

A.M. Prudnik

Associate Professor of Engineering Psychology and Ergonomics Department of BSUIR, Candidate of Technical Sciences, Associate Professor

Annotation: The article explores the potential of virtual reality (VR) in the field of shooting training, especially in the context of pre-prescription training and training of military-patriotic values among young people. The development of scenarios for shooting range game simulators in VR is considered, which are effective tools for improving accuracy, reaction and shooting skills. Various types of shooting ranges are discussed, including stationary and moving targets, as well as gameplay control scripts. The study focuses on the potential of virtual reality in improving the quality of pre-prescription training and improving the effectiveness of training in the field of shooting and weapons.

Keywords: VR, Unity, virtual reality, shooting training, shooting range simulators, game mechanics, pre-prescription training, educational technologies, virtual simulators.

УДК 004.722:004.75:004.67

СЕТЕВАЯ ИНФРАСТРУКТУРА ДЛЯ РАБОТЫ С BIG DATA



М.В. Романюк

*И.о. начальника отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики, магистрант кафедры ПИКС
romanuk@bsuir.by*



Е.А. Лещенко

*Инженер-программист отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики, магистрант кафедры ПИКС
e.leshchenko@bsuir.by*



А.Н. Марков

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

М.В. Романюк

Окончил Белорусский государственный университет информатики и радиоэлектроники. Инженер-программист, и.о. начальника отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики БГУИР, магистрант кафедры проектирования информационно-компьютерных систем БГУИР.

Е.А. Лещенко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Инженер-программист, отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики БГУИР, магистрант кафедры проектирования информационно-компьютерных систем БГУИР.

А.Н. Марков

Окончил Белорусский государственный университет информатики и радиоэлектроники. Заместитель начальника Центра информатизации и инновационных разработок БГУИР, старший преподаватель. Область научных интересов связана с облачными вычислениями, распределенными вычислительными системами, балансировкой нагрузки вычислительных систем.

Аннотация. В данной работе рассмотрены популярные архитектуры компьютерных сетей, а также требования к компьютерным сетям, в которых выполняется передача *Big Data*.

Показано, что наиболее подходящей является архитектура *Fabric (Spine-Leaf)*, которая обеспечивает предсказуемость скорости задержек при передаче данных, что положительно сказывается на поддержании высокой синхронности работы приложений на различных хостах в центре обработки данных.

Ключевые слова: spine-leaf, сеть fabric, передача больших данных

Введение. В настоящее время крупные организации, включая учебные заведения, по всему миру вынуждены обрабатывать большие объемы данных, передавать эти данные по сети между серверами, обеспечивая при этом высокую степень отказоустойчивости ИТ-инфраструктуры.

Многие приложения для работы с большими данными работают в режиме реального времени. Следовательно, эти приложения должны создавать, хранить и обрабатывать большое количество информации, что приводит к увеличению объема передачи и

нагрузке на сеть. Таким образом, к важной области исследований в данном случае можно отнести оптимизацию топологии сети.

Требования к сетевой инфраструктуре. Обработка *Big Data* предъявляет высокие требования к производительности сетевой инфраструктуры, что означает, что сеть должна быть устойчивой, согласованной и обладать определенной степенью осведомленности о приложениях. Идеальная архитектура сети больших данных должна разрабатываться с учетом распределенной архитектуры, чтобы обеспечить доступность распределенных ресурсов при одновременной параллельной работе над одной задачей [1].

Таким образом, можно сказать, что основными задачами при организации сети для обработки *Big Data* являются:

1. Отказоустойчивость сети включает в себя идею возможности использования различных маршрутов при попытке доступа к ресурсам и способности быстро распознавать проблемы и переключаться на альтернативный путь.

2. Снижение перегрузки сети. Под перегрузкой сети понимается проблема, при которой линии связи не могут передать необходимый объем данных, что вызывает задержки в очередях, проблемы с потерянными пакетами и повторной передачей.

3. Согласованность производительности сети. Важным является поддержание высокой синхронности. Различные задания приложений должны выполняться параллельно на различных хостах, поэтому любое значительное снижение производительности сети может спровоцировать рассинхронизацию и сбои в результатах.

4. Возможность масштабирования сети должна включать в себя как возможность легкого расширения полосы пропускания, так и возможность расширения портовой емкости для подключения дополнительных серверов.

Требования к потокам данных. Традиционно сетевая инфраструктура создавалась в виде трехуровневой иерархической модели и предполагала продвижение трафика «Север-Юг», что означает, что означает движение трафика из-за пределов локальной сети к определенному серверу, на котором хранится информация, а также обратный путь трафика.

С внедрением *Big Data* в сетях схема трафика изменилась. Данные больше не создаются только внешними источниками, а генерируются множеством связанных устройств внутри центра обработки данных. Этот тип трафика называется межмашинным/виртуальным или трафиком «Запад-Восток».

Традиционная сетевая инфраструктура. Большая часть традиционных современных сетевых сред основана на иерархической структуре, где в нижней части дерева расположен уровень доступа (*access, rack*), к которому подключаются хосты (серверы). Средний уровень – это уровень распределения (агрегации), который объединяет коммутаторы уровня доступа, позволяет выполнить настройки управления и маршрутизации и делегирует *Uplink* на более высокий уровень. Верхний уровень (ядро) отвечает за маршрутизацию и предоставление доступа к услугам за пределами локальной вычислительной сети (рисунок 1). При использовании такой архитектуры узкими местами чаще всего являются ядро, которое может быть перегружено трафиком Север-Юг от пользовательских приложений, а также перегруженными могут быть восходящие и нисходящие каналы от уровня доступа к уровню агрегации, которые нагружаются как при пересылке трафика Север-Юг, так и при передаче в пределах центра обработки данных (Запад-Восток), а также из-за блокировки избыточных каналов (при условии использования протокола *STP*). Ввиду всего сказанного можно заметить, что иерархическая модель не способна обеспечить одинаковые и прогнозируемые задержки при передаче данных, что требуется при работе с *Big Data* [23].

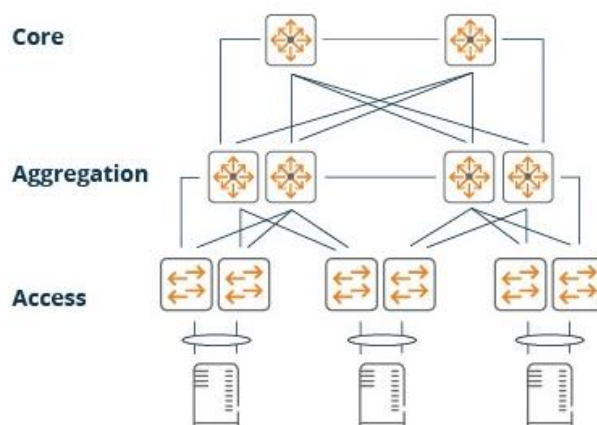


Рисунок 1. Схема традиционной трехуровневой архитектуры Core-Aggregation-Access

Современная архитектура Fabric (Spine-Leaf). Архитектура *Spine-Leaf* спроектирована специально для обработки больших объемов трафика, который передается в пределах центра обработки данных (Запад-Восток). *Leaf*-коммутаторы образуют уровень доступа. Каждый *Leaf* коммутатор связан с каждым *Spine*-коммутатором. Таким образом архитектура *Spine-Leaf* гарантирует, что коммутаторы уровня доступа находятся на расстоянии не более одного хопа друг от друга, что минимизирует задержки и вероятность возникновения узких мест между коммутаторами уровня доступа.

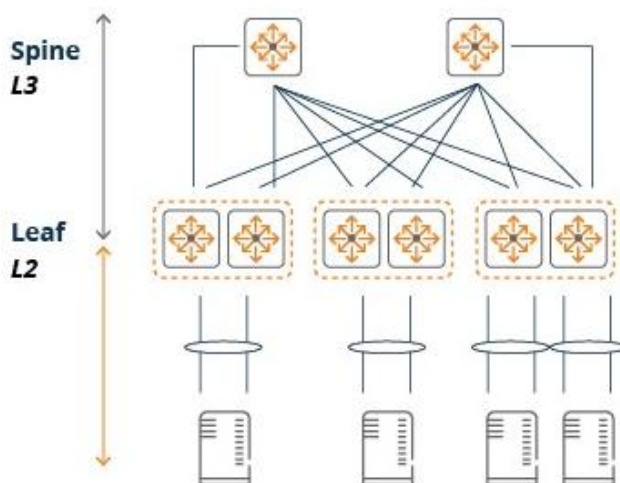


Рисунок 2. Схема сетевой архитектуры Fabric (Spine-Leaf)

Архитектура «*Spine-Leaf*» может быть *L2* или *L3*, что означает, что каналы между уровнем «*Leaf-Spine*» могут быть либо коммутируемыми, либо маршрутизируемыми. В любом случае все каналы являются пересылающими, т. е. ни один из них не блокируется, поскольку *STP* заменяется другими протоколами, например *ECMP* [2].

При этом сети *Spine-Leaf* обладают и недостатком: основной недостаток происходит от использования маршрутизации *Layer 3*, который не дает развертывать виртуальные сети *VLAN* во всей сети. Сети *VLAN* в архитектуре *Leaf-Spine* локализируются в каждом из *Leaf*-коммутаторов и другим *Leaf*-коммутаторам недоступны. Это может создавать проблемы при обеспечении мобильности гостевой виртуальной машины в пределах дата-центра.

Для решения этой проблемы, может быть использовано решение программно-конфигурируемых сетей *SDN* (*Software Defined Networking*), которое создает виртуальный

уровень *L2* поверх сети *Spine-Leaf*. При этом виртуальные серверы могут свободно перемещаться внутри среды дата-центра, без ухудшения эффективности трафика Запад-восток, масштабируемости и стабильности топологии сети *Leaf-Spine* [3].

Как альтернатива маршрутизации *Layer 3*, в топологии *Leaf-Spine* могут использоваться и другие протоколы, такие как *TRILL* (*Transparent Interconnection of Lots of Links*) или *SPB* (*Shortest Path Bridging*) [4].

Разделение уровней управления и данных позволяет динамически настраивать сетевые конфигурации начиная от расстановки приоритетов трафика и заканчивая созданием динамических правил маршрутизации для конкретного приложения в соответствии с быстро меняющимися потребностями.

Заключение. Таким образом, рассмотрение потребностей сетей *Big Data* и распространенных сетевых топологий показало, что идеальным вариантом для реализации сети для передачи больших данных является использование топологии *Fabric (Spine-Leaf)*, которая имеет множество преимуществ по сравнению с иерархической архитектурой: повышенная пропускная способность, производительность, балансировка нагрузки, масштабируемость, возможность эффективной интеграции с *SDN* для еще более точного контроля над трафиком.

Список литературы

[1] Big Data: архитектура w/vs инфраструктура: [Электронный ресурс]. URL: https://habr.com/ru/companies/ru_mts/articles/531818/. (Дата обращения: 12.02.2024).

[2] Banks E. Data center network design moves from tree to leaf: [Электронный ресурс]. URL: <https://www.techtarget.com/searchdatacenter/feature/Data-center-network-design-moves-from-tree-to-leaf>. (Дата обращения: 13.02.2024).

[3] Deep P. Big Data Analytics and Software Defined Networks: [Электронный ресурс]. URL: <https://medium.com/@20bt04032/big-data-analytics-and-software-defined-networks-a01a131d889>. (Дата обращения: 12.02.2024).

[4] Шалагинов А. Сетевая топология Leaf-Spine: [Электронный ресурс]. URL: <https://shalaginov.com/2016/03/26/сетевая-топология-leaf-spine/>. (Дата обращения: 10.02.2024).

Авторский вклад

Авторы внесли равноценный вклад.

FABRIC NETWORK INFRASTRUCTURE FOR BIG DATA

M.V. Romaniuk

Acting Head of the Network Technologies Department of the Center of Informatization and Innovation Elaborations, Assistant of the Department of Computer Science of BSUIR, Master's student of the Department of ICSD of BSUIR

E.A. Leshchenko

Software Engineer of the Network Technology Department of the Center of Informatization and Innovation Elaborations, Assistant of the Department of Computer Science of BSUIR, Master's student of the Department of ICSD of BSUIR

A.N. Markov

Senior lecturer, Master of Technical Sciences, Deputy head of the Center of Informatization and Innovation Elaborations

Abstract. This article describes popular computer network architectures, as well as the requirements for computer networks in which Big Data transmission is performed.

It is shown that the most suitable architecture is Fabric (Spine-Leaf), which provides predictability of latency rate during data transmission, which positively affects the maintenance of high synchronization of applications on different hosts in the data center.

Keywords: spine-leaf, fabric network, big data transmission

УДК 004.021:004.75

РАСПРЕДЕЛЕННАЯ СИСТЕМА ПОТОКОВОЙ ОБРАБОТКИ ДАННЫХ ДЛЯ ЗАДАЧ РАСПОЗНАВАНИЯ РЕЧИ



К. Жакшылык

Магистрант кафедры ИТАС, факультета информационных технологий и управления БГУИР
kuanysh.zhk@gmail.com



В. А. Захарьев

Доцент кафедры систем управления БГУИР,
кандидат технических наук
zaharyev@bsuir.by

К. Жакшылык

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой методов и алгоритмов построения высоконагруженных распределенных систем для потоковой обработки больших объемов данных, построение аналитических платформ на базе хранилищ данных.

В.А.Захарьев

Кандидат технических наук, доцент кафедры систем управления БГУИР. Имеет более 50 научных публикаций. Научные интересы лежат в области методов цифровой обработки сигналов и машинного обучения, а также разработки человеко-машинных интерфейсов на основе синтеза и распознавания речи.

Аннотация. Представлен обзор архитектурных решений для распределенных систем потоковой обработки данных, предназначенных для построения современных сервисов распознавания речи на основе глубоких нейросетевых моделей. Рассмотрены основные компоненты таких систем, включая слои хранения и обработки данных, а также особенности их программной реализации. Особое внимание уделено использованию «*Apache Kafka*» в качестве брокера сообщений для обеспечения эффективной передачи данных между компонентами системы. Показаны особенности применения «*Mflow*» для развертывания модели распознавания речи «*Whisper*», что обеспечивает удобное управление жизненным циклом модели и ее метаданными. Рассмотрен процесс развертывания приложения в рамках концепции микросервисной архитектуры на базе системы управления вычислительным кластером «*Kubernetes*», предоставляющей широкие возможности масштабирования вычислительных ресурсов. Полученные результаты позволяют определить ключевые характеристики распределенных систем, влияющие на эффективность работы моделей распознавания речи, работающих в реальном времени

Ключевые слова: распределенные системы, распознавание речи, микросервисная архитектура, брокер сообщений «*Apache Kafka*», система управления вычислительным кластером «*Kubernetes*», модель распознавания речь-в-текст «*Whisper*», система развертывания нейросетевых моделей «*MLflow*».

Введение. В контексте широкого распространения цифровых коммуникаций и проникновения интеллектуальных устройств в различные сферы жизни запрос на создание эффективных и точных системы распознавания речи достиг существенного уровня. В условиях возрастающих потребностей использования речевых технологий как разработчики, так и пользователи, сталкиваются с необходимостью обеспечения высокой производительности и масштабируемости моделей распознавания речи. Традиционные подходы к этой задаче часто сталкиваются с проблемами, связанными с огромным объемом и динамическим характером аудиоданных в реальном времени. В ответ на эти

вызовы интеграция распределенных вычислений и обработки потоков данных стала привлекательной парадигмой, которая обещает улучшить скорость, точность и масштабируемость систем распознавания речи.

Современные речевые системы и сервисы, обладающие хорошими показателями качества, из-за высокой степени сложности и ресурсоемкости используемых моделей, имеют распределенную архитектуру. Следовательно, возникает необходимость в развитии методов проектирования, а также, в создании архитектур и алгоритмов распределённых систем, позволяющих обеспечить: повышение отказоустойчивости подобных систем, увеличение скорости доступа к моделям распознавания речи, находящимся на удалённых ресурсах, возможность их масштабирования в зависимости от нагрузки.

1 Системы потоковой обработки данных.

Рядом специалистов [1,2] были рассмотрены основные принципы построения систем потоковой обработки данных, предложен архитектурный вариант реализации распределенной системы потоковой обработки данных для задач распознавания речи на базе модели *Whisper* с *Apache Kafka* и *Kubernetes*.

Традиционный метод конфигурации инфраструктуры для обработки больших данных часто включает в себя развертывание «*Hadoop*»-кластера и настройку соответствующих инструментов для построения платформы обработки данных. Однако данный подход сталкивается с рядом ограничений, таких как отсутствие возможности четкого разделения между уровнями хранения и вычислений, сложности в масштабировании и недостаточная изоляция сред для различных приложений. Даже при предоставлении «*Hadoop*» в качестве сервиса («*Hadoop-as-a-Service*») облачным провайдером, подход остается схожим с традиционным развертыванием на собственном оборудовании.

Однако существует альтернативный, основанный на облачных технологиях: подход к обработке больших данных, известный как «*Cloud-Native*». Этот подход не только позволяет преодолеть упомянутые ограничения, но и предоставляет дополнительные преимущества, основанные на возможностях облачных технологий. Для реализации такого подхода в качестве инфраструктуры для построения системы используется платформа управления вычислительными ресурсами «*Kubernetes*», которая совместно с другими компонентами, обеспечивает большую гибкую и масштабируемую среду для обработки и анализа данных. Далее будут рассмотрены основные концепции для построения систем потоковой обработки данных.

Потоковая обработка данных – это метод обработки информации, при котором данные анализируются по мере их непрерывного поступления, обеспечивая оперативное реагирование и принятие решений. В сравнении с традиционной пакетной обработкой, где данные группируются и обрабатываются в определенные интервалы времени (пакеты), потоковая обработка работает в режиме реального времени, что позволяет извлекать ценную информацию из потока данных немедленно. Далее представлен алгоритм и основные принципы потоковой обработки данных.

Захват данных:

Процесс начинается с извлечения информации из различных источников данных, таких как сенсоры, приложения и веб-сервисы.

Главная цель этого этапа – обеспечить эффективный и надежный сбор данных для последующей обработки.

Агрегация и преобразование:

После захвата данных они проходят через этап агрегации и преобразования. На этом этапе данные могут быть отфильтрованы, преобразованы в нужный формат или агрегированы для упрощения последующей обработки.

Обработка:

Данные, уже агрегированные и преобразованные, направляются на обработку. Здесь применяются различные алгоритмы и логика для анализа данных, выявления шаблонов, трендов и аномалий.

Принятие решений:

После обработки данных система может автоматически принимать решения на основе полученных результатов. Это может включать в себя автоматические действия или предоставление информации для принятия решений человеком.

Действие и хранение:

В зависимости от результатов обработки, система может выполнять дополнительные действия. Это может включать отправку уведомлений, запись данных в хранилище, архивирование информации или внесение изменений в рабочие процессы.

2 Анализ моделей потоковой обработки данных.

Перед тем как переходить непосредственно к рассмотрению архитектуры распределённой системы необходимо определиться с выбором базовой модели потоковой обработки данных. Можно выделить две основные такие модели: точка-точка и микросервисная модель [6].

Модель «точка-точка» в архитектуре потоковой обработки данных представляет собой подход, при котором данные передаются напрямую от одного источника к одному или нескольким конечным потребителям.

Микросервисная модель в контексте потоковой обработки данных представляет собой подход, при котором сложные системы обработки данных разделяются на небольшие, автономные компоненты, называемые микросервисами.

Произведём сравнительный анализ моделей построения распределённой системы используя метод их качественного сопоставления, который подразумевает систематическое изучение и анализ характеристик двух или более подходов, моделей или явлений с целью выявления их сходств, различий, преимуществ и недостатков. Результаты качественного анализа моделей представлены в табл. 1.

Таблица 1. Результаты качественного анализа моделей потоковой обработки данных.

Модели потоковой обработки данных	Преимущества	Недостатки
Точка-точка	Низкая задержка: поскольку данные передаются непосредственно от источника к потребителю без промежуточных этапов обработки, задержка минимальна, что особенно важно для приложений, требующих реакции в реальном времени.	Отсутствие гибкости: Такая модель не обеспечивает гибкости в управлении данными и их обработке на промежуточных этапах. Возникают сложности при добавлении новых источников или потребителей данных.

Продолжение таблицы 1

Модели потоковой обработки данных	Преимущества	Недостатки
	<p>Простота реализации: Эта модель является относительно простой в реализации и управлении, так как не требует сложной логики маршрутизации и обработки данных.</p>	<p>Отсутствие возможности анализа и преобразования данных: поскольку данные передаются непосредственно, без промежуточной обработки, они не могут быть анализированы или преобразованы на этапе передачи.</p>
	<p>Эффективное использование ресурсов: поскольку данные напрямую доставляются от источника к потребителю, нет необходимости хранить данные на промежуточных узлах, что способствует экономии ресурсов.</p>	<p>Ограниченная масштабируемость: при увеличении количества источников и потребителей модель «точка-точка» может столкнуться с ограничениями масштабируемости из-за необходимости поддерживать прямые соединения между каждой парой источник-потребитель</p>
Микросервисная модель	<p>Гибкость: Модель «микросервисы» позволяет гибко масштабировать и развивать систему путем добавления, удаления или изменения микросервисов.</p>	<p>Мониторинг и отладка: Следить за работой нескольких микросервисов и находить ошибки может быть более сложно по сравнению с другими моделями</p>
	<p>Изоляция ошибок: Изоляция микросервисов позволяет ограничить влияние ошибок в одном компоненте на другие части системы, обеспечивая надежность и стабильность работы.</p>	
	<p>Независимая разработка: Команды разработчиков могут работать над отдельными микросервисами независимо друг от друга, ускоряя процесс разработки.</p>	

Окончание таблицы 1

Модели потоковой обработки данных	Преимущества	Недостатки
	Технологическое разнообразие: Разные микросервисы могут быть реализованы с использованием различных технологий, что позволяет выбирать наилучшие инструменты для каждой задачи.	

Исходя из результатов анализа, представленных выше, можно сделать выводы что микросервисная модель представляет собой привлекательную и грамотную выборку в контексте потоковой обработки данных, при условии ее правильной реализации.

3 Архитектура распределённой системы потоковой обработки.

Типовая архитектура потоковой обработки данных выделяет два ключевых слоя, необходимых для организации системы:

- слой хранения (*storage layer – SL*);
- слой обработки (*processing layer – PL*).

Слой хранения должен обеспечивать возможность организации высокой пропускной способности для операций ввода/вывода для больших потоков данных. Слой обработки отвечает за потребление данных, поступающих из слоя хранения, обработку этих данных и оповещение *SL* об обновлении данных [9].

Типовая двухслойная архитектура, состоящая из слоя хранения и слоя обработки, имеет свои ограничения, особенно в контексте распределенных систем и обработки больших объемов данных. Например, она может столкнуться с проблемами масштабирования, управления ресурсами и обеспечения отказоустойчивости в условиях высокой нагрузки. Кроме того, она может оказаться недостаточно гибкой для обработки разнообразных источников данных и реализации сложных алгоритмов обработки.

Для организации распределенной обработки потоков данных на множестве независимых вычислительных устройств, требуется некоторое расширение представленной модели. Исходя из этих принципов и из принципов построения данных систем, была составлена архитектура системы, состоящая из пяти слоев (см. рис. 1) [10]:

1 Слой потребления потока данных, включает операции по получению потоков данных из первоначальных источников и передаче в систему обработки или хранения данных [8]. Это могут быть различные *IoT* – решения, датчики, *WebSocket* соединения, системы логирования, брокеры сообщений, очереди и т.п.

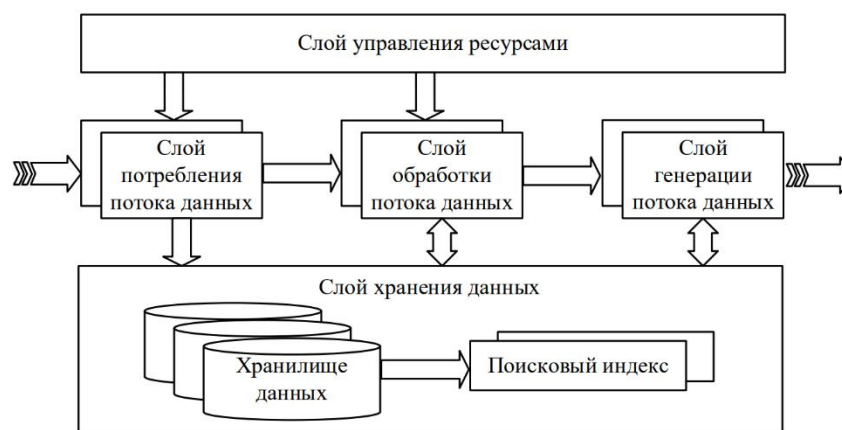


Рисунок 1. Архитектура потоковой системы обработки данных

2 Слой обработки потока данных, это слой, на котором выполняются приложения для обработки потоковых данных. На нем могут размещаться как отдельные приложения обработки данных, так и платформы обработки потоков данных (*Data Stream Processing Engines – DSPE*) [1]. Например, *Azure Databricks*, единая, открытая платформа аналитики для создания, развертывания, совместного использования и обслуживания корпоративных данных, аналитики и решений искусственного интеллекта в масштабе.

3 Слой хранения данных, обеспечивает хранение сообщений, промежуточных и финальных результатов обработки, выявленных паттернов и информации из потока данных [2]. Различные решения могут обеспечить поддержку такого слоя, включая хранилища «ключ-значение», реляционные и *NoSQL* базы данных, базы данных в памяти [3].

4 Слой управления ресурсами отвечает за организацию работы и взаимодействие вычислительных узлов, узлов хранения, а также за управление временем жизни данных ресурсов (включая выделение новых и освобождение занятых ресурсов) для поддержки обработки больших объемов потоков данных [4]. В качестве примера может выступать служба для координации распределённых систем, организованная на основе резидентной базы данных категории «ключ – значение» *Apache Zookeeper*, *Kubernetes* для оркестровки контейнеризированных приложений.

5 Слой генерации потока данных отвечает за формирование результирующего потока данных и его отправку последующим потребителям. В качестве таких потребителей могут выступать сторонние приложения, другие системы обработки данных, системы визуализации и мониторинга и др. [3].

4 Основные компоненты архитектуры и особенности их реализации.

В качестве основного компонента предлагаемой архитектуры слоя потребления данных в работе предлагается использовать брокер сообщений «*Apache Kafka*», представляющий собой одновременно серверное программное обеспечение, реализующее функции распределённого брокера (очереди) сообщений, а также соответствующий фреймворк для различных языков программирования, позволяющих подключаться к серверу и передавать сообщения. Работа с *Apache Kafka* основана на использовании следующих основных программных абстракций и компонентов:

- брокеры;
- продюсеры;
- потребители.

Брокер – это экземпляр сервера «*Apache Kafka*» (также известный как узел *Kafka*), который размещает именованные потоки записей, называемые темами. Брокер принимает сообщения от продюсеров и сохраняет их в темы. В свою очередь, он позволяет потребителям извлекать сообщения из темы.

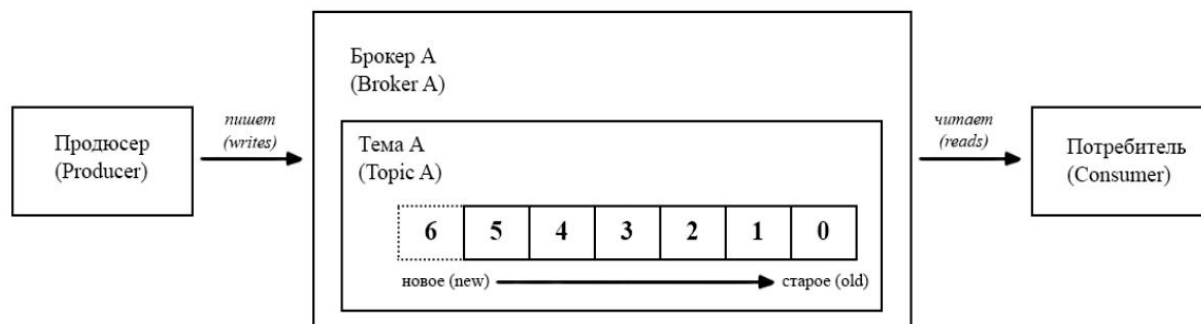


Рисунок 2. Архитектура брокера сообщений на основе «*Apache Kafka*»

В случае использования базового варианта архитектуры кластера «*Kafka*» может функционировать с одним брокером сообщений. Такой вариант использования приводит к появлению в системе единой точки отказа, что может нарушить связность компонентов в системе: препятствовать продюсерам отправлять сообщения, а потребителям получать их. Чтобы обеспечить доступность, нам нужны запасные брокеры для резервного копирования.

Компонент «*Kafka*» является приложением с состоянием, что означает, что ему нужно сохранять данные и поддерживать свое состояние. Развертывание приложения с состоянием представляет собой довольно сложную задачу. Для решения данной проблемы в рамках предложенной архитектуры воспользуемся системой управления вычислительными ресурсами «*Kubernetes*», которая эффективно позволяет справиться с задачей создания и управления группой персистентных (постоянных) томов.

Из представленной схемы на рисунке 3 можем выделить следующие ресурсы:

1 Кластер *Kafka* с тремя брокерами. Основное преимущество такого кластера *Kafka* заключается в его доступности. В случае если один из брокеров сообщений выйдет из строя, *Kafka* всё равно способна обрабатывать запросы своих клиентов (как потребителей, так и продюсеров).

2 В сочетании с *StatefulSet* мы используем так называемый «безымянный» сервис (*headless service*), чтобы у каждого пода *Kubernetes* были свои собственные *DNS*-имена. Это необходимо, так как *IP*-адреса подов могут изменяться, и нам нужно обеспечить возможность правильного обращения к подам.

3 Каждый из подов брокеров использует собственное постоянное хранилище на базе механизма *Kubernetes* «*persistent volume*».

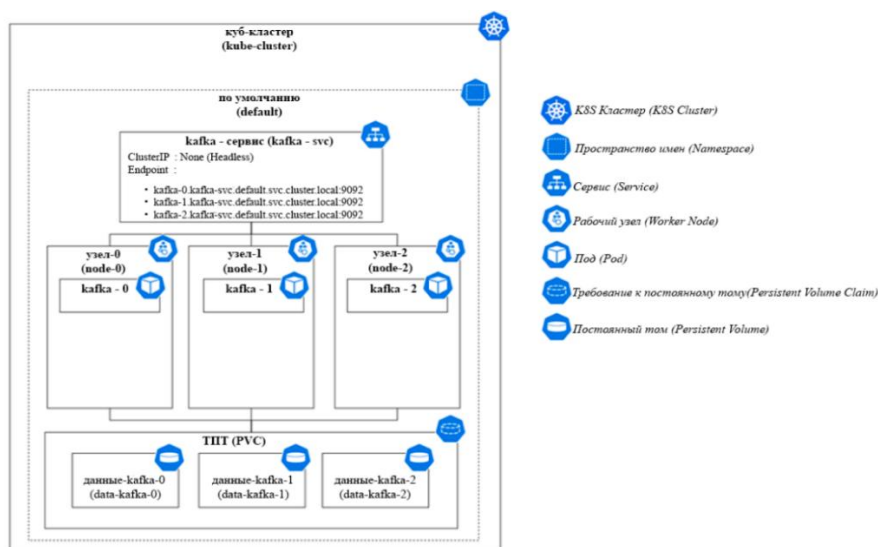


Рисунок 3. Диаграмма развертывания брокера сообщений «Apache Kafka», по средствам системы управления вычислительными ресурсами «Kubernetes»

Далее представлен Манифест для создания мультиброкерского кластера *Kafka* в *Kubernetes*:

```

apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: kafka
  labels:
    app: kafka-app
spec:
  serviceName: kafka-svc
  replicas: 3
  selector:
    matchLabels:
      app: kafka-app
  template:
    metadata:
      labels:
        app: kafka-app
    spec:
      containers:
        - name: kafka-container
          image: doughgle/kafka-kraft
          ports:
            - containerPort: 9092
            - containerPort: 9093
      env:
        - name: REPLICAS
          value: «3»
        - name: SERVICE
          value: kafka-svc
        - name: NAMESPACE

```

```
value: default
- name: SHARE_DIR
value: /mnt/kafka
- name: CLUSTER_ID
value: kb-WV7dt8Au9ixOk4YbRm5tL
- name: DEFAULT_REPLICATION_FACTOR
value: «3»
- name: DEFAULT_MIN_INSYNC_REPLICAS
value: «2»
volumeMounts:
- name: data
mountPath: /mnt/kafka
volumeClaimTemplates:
- metadata:
name: data
spec:
accessModes:
- «ReadWriteOnce»
resources:
requests:
storage: «1Gi»
```

Обработка входящих потоков моделью распознавания речи. После конфигурации Kubernetes кластера и Apache Kafka, необходимо настроить платформу по обработке входящих потоков и дальнейшей процесс распознавания речи. Ниже приведен блок кода для потребителя в Apache Kafka.

```
from kafka import KafkaProducer
topic = "audio-stream"

def publish_audio():
    """
    Отправляет аудиопоток с микрофона в указанную тему Kafka.
    """
    # Начало работы продюсера
    producer = KafkaProducer(bootstrap_servers='localhost:9092')
    # Настройка аудиопотока
    CHUNK = 1024
    FORMAT = pyaudio.paInt16
    CHANNELS = 1
    RATE = 44100
    audio = pyaudio.PyAudio()
    stream = audio.open(format=FORMAT, channels=CHANNELS,
                        rate=RATE, input=True,
                        frames_per_buffer=CHUNK)
    print("Отправка аудиопотока...")

    try:
        while True:
            data = stream.read(CHUNK)
            producer.send(topic, data)
```

```
time.sleep(0.1) # Спящий режим для управления скоростью потоковой передачи
```

```
except KeyboardInterrupt:  
    print("\nЗавершение работы.")  
    stream.stop_stream()  
    stream.close()  
    audio.terminate()  
    producer.close()  
    sys.exit(1)
```

```
if __name__ == '__main__':  
    """
```

```
    Продюсер будет отправлять аудиопоток с микрофона на сервер Kafka.  
    """
```

```
    publish_audio()
```

Эффективное развертывание моделей машинного обучения в производственной среде является критическим фактором, поскольку любая задержка может привести к устареванию данных и утрате воспроизводимости экспериментов. Однако часто процесс передачи моделей от специалистов по данным к инженерам данных требует значительного времени, что может препятствовать оперативному внедрению новых моделей. Решить эту проблему помогает практика *MLOps*, которая стандартизирует процесс разработки моделей машинного обучения.

Одним из ключевых инструментов в рамках практики *MLOps* является *MLflow* – платформа для машинного обучения и *Data Science*. *MLflow* предоставляет средства для эффективного управления жизненным циклом моделей, включая наблюдение экспериментов, моделей и артефактов. Кроме того, *MLflow* обеспечивает возможность оперативного вывода моделей в производство и предоставления их в виде сервиса всего за несколько минут.

В качестве модели распознавания речи будет использоваться *Whisper* от лаборатории *OpenAI*. *Whisper*, разработанный *OpenAI*, представляет собой универсальную модель *ASR*, обученную для точного преобразования речи в текст. Его особенность заключается в том, что она была обучена на 680 000 часов многоязычных и многозадачных контролируемых данных, собранных из Интернета [5].

Преимущества использования модели *Whisper* с *MLflow*:

Отслеживание экспериментов: облегчает отслеживание конфигураций модели и ее производительности для достижения оптимальных результатов.

Управление моделями: централизует различные версии моделей *Whisper*, улучшая доступность.

Воспроизводимость: обеспечивает согласованность транскрипций путем отслеживания всех компонентов, необходимых для воспроизведения поведения модели.

Развертывание: упрощает развертывание моделей *Whisper* в различных производственных средах, обеспечивая эффективное применение.

Перед тем как приступить к процессу аудиотранскрипции с помощью модели *Whisper* необходимо выполнить несколько подготовительных шагов, чтобы обеспечить плавное и эффективное преобразование аудио в текст.

Получение аудио. Первым шагом является получение аудиофайла для работы. Мы используем общедоступный аудиофайл от *NASA*. Этот образец аудио обеспечивает практический пример для демонстрации возможностей транскрипции *Whisper*.

Инициализация модели и конвейера. Мы загружаем модель *Whisper*, а также ее токенизатор и извлекатель признаков из библиотеки *Transformers*. Эти компоненты необходимы для обработки аудиоданных и преобразования их в формат, понятный модели *Whisper* для транскрипции. Затем мы создаем конвейер транскрипции, используя модель *Whisper*. Этот конвейер упрощает процесс подачи аудиоданных в модель и получения транскрипции.

Настройка среды *MLflow*. Помимо настройки модели и аудиоданных, мы инициализируем среду *MLflow*. *MLflow* используется для отслеживания и управления нашими экспериментами, предлагая организованный способ документирования процесса транскрипции и результатов [7].

Нижеприведенный кодовый блок охватывает эти начальные шаги настройки, обеспечивая основу для нашей задачи аудиотранскрипции с моделью *Whisper*. А также создает *KafkaConsumer* для получения данных и сообщений, созданных *KafkaProducer*:

```
import json
from kafka import KafkaConsumer
import mlflow
import requests
import transformers

# Функция для обработки аудио и распознавания речи с использованием Whisper и
MLflow
def process_audio(audio):
    # Установка задачи, которую будет использовать наша реализация конвейера
    task = "автоматическое-распознавание-речи"

    # Определение экземпляра модели
    architecture = "openai/whisper-large-v3"

    # Загрузка компонентов и необходимой конфигурации для Whisper ASR из Hugging
Face Hub
    model = transformers.WhisperForConditionalGeneration.from_pretrained(architecture)
    tokenizer = transformers.WhisperTokenizer.from_pretrained(architecture)
    feature_extractor =
transformers.WhisperFeatureExtractor.from_pretrained(architecture)
    model.generation_config.alignment_heads = [[2, 2], [3, 0], [3, 2], [3, 3], [3, 4], [3, 5]]

    # Создание конвейера для ASR с использованием модели Whisper
    audio_transcription_pipeline = transformers.pipeline(
        task=task, model=model, tokenizer=tokenizer, feature_extractor=feature_extractor
    )

    # Проверка того, что наш конвейер способен обработать аудиофайл и провести
транскрибирование
    transcription = audio_transcription_pipeline(audio)
    return transcription

# Настройки потребителя Kafka
topic = "audio-stream"
bootstrap_servers = 'localhost:9092'

# Настройка Kafka потребителя
```

```
consumer = KafkaConsumer(topic,  
                           bootstrap_servers=bootstrap_servers,  
                           value_deserializer=lambda m: m)  
  
# Настройка MLflow  
mlflow.set_tracking_uri("http://mlflow-server:5000")  
mlflow.set_experiment("эксперимент-распознавания-речи")  
  
# Обработка аудио и распознавание речи для каждого сообщения из Kafka  
for message in consumer:  
    audio_data = message.value # Получение аудиоданных из Kafka  
    # Обработка аудио с использованием Whisper и MLflow  
    with mlflow.start_run():  
        transcription = process_audio(audio_data)  
        # Логирование результатов в MLflow  
        mlflow.log_text("transcription", transcription["text"])
```

Заключение. В статье были рассмотрены существенные аспекты создания распределённых систем потоковой обработки информации для систем распознавания речи. Представлены существующие и перспективные методы построения систем потоковой обработки, включая архитектурный вариант с использованием *Apache Kafka* и *Kubernetes* для задач распознавания речи. Выполнен сравнительный качественный анализ моделей «точка-точка» и микросервисной архитектуры, их преимуществ и недостатков в контексте распределённых систем. Предложена архитектура распределённой системы потоковой обработки, включая слои хранения и обработки данных. Показаны их основные ограничения и возможности расширения для дальнейшего улучшения. Детально рассмотрены ключевые компоненты архитектуры и особенности их реализации, такие как: брокер сообщений *Apache Kafka* и система управления вычислительными ресурсами *Kubernetes*, и их роль в обеспечении масштабируемости и отказоустойчивости системы. Приведены программные артефакты описания процесса конфигурации и использования модели распознавания речи *Whisper* для обработки и транскрибирования аудиоданных.

Разработанная система демонстрирует, как использование *Apache Kafka*, *Kubernetes* и *MLflow* вместе с моделью распознавания речи *Whisper* от *OpenAI* улучшает масштабируемость, отказоустойчивость и точность систем распознавания речи. Система применима в различных сферах, не только непосредственно для распознавания речи, но и, например, для зада автоматизации обработки обращений и создание автоматических субтитров. Отмечено, что в процессе реализации важно учитывать потребности в вычислительных ресурсах, безопасность данных и предотвращение сбоев. Дальнейшее усовершенствование системы может включать оптимизацию архитектуры, улучшение обработки данных и интеграцию с другими сервисами. Результаты исследования подтверждают значимость развития распределённых систем обработки данных для построения эффективного высоконагруженных распределённых сервисов обработки пользовательских данных.

Список литературы

[1] Alaasam A.B.A. et al. Analytic Study of Containerizing Stateful Stream Processing as Microservice to Support Digital Twins in Fog Computing // Programming and Computer Software. 2020. Vol. 46, no. 8. P. 511–525. DOI:10.1134/S0361768820080083

[2] Isah H. et al. A Survey of Distributed Data Stream Processing Frameworks // IEEE Access. IEEE, 2019. Vol. 7, no. October. P. 154300–154316. DOI:10.1109/ACCESS.2019.2946884. Meehan J., Zdonik S. Data Ingestion for the Connected World // Cidr. 2017.

- [3] Gualtieri M., Yuhanna N. The forrester wave: Big data streaming analytics, Q1 2016 // Forrester research. Cambridge, MA, USA, 2016. 15 p
- [4] Henning S., Hasselbring W. Theodolite: Scalability Benchmarking of Distributed Stream Processing Engines in Microservice Architectures // Big Data Research. 2021. Vol. 25. DOI:10.1016/j.bdr.2021.100209.
- [5] Real-time full-text search with Luwak and Samza - Confluent [Electronic resource]. URL: <https://www.confluent.io/blog/real-time-full-text-search-with-luwak-and-samza/> (accessed: 27.04.2021).
- [6] Dias de Assunção M., da Silva Veith A., Buyya R. Distributed data stream processing and edge computing: A survey on resource elasticity and future directions // Journal of Network and Computer Applications. Elsevier Ltd, 2018. Vol. 103. P. 1–17. DOI:10.1016/j.jnca.2017.12.001.
- [7] Kleppmann, Martin. Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems. « O'Reilly Media, Inc.», 2017.
- [8] Tanenbaum, Andrew S. Distributed systems principles and paradigms. 2007.
- [9] Vitillo, Roberto. Understanding Distributed Systems: What every developer should know about large distributed applications. Roberto Vitillo, 2022.
- [10] Joshi, Unmesh. "Patterns of distributed systems." martinowler. com (2020).

Авторский вклад

Жаксылык Куаныш – руководство исследованием по разработке распределенной системы потоковой обработки данных для задач распознавания речи. Тестирование распределенной системы, развертывание кластера Kubernetes, Apache Kafka, оценка результатов.

Захарьев Вадим Анатольевич – постановка задачи исследования, описание принципа работы платформы MLflow и модели распознавания речи Whisper, формирование структуры статьи.

DISTRIBUTED STREAM DATA PROCESSING SYSTEM FOR SPEECH RECOGNITION TASKS

K. Zhaksylyk

Master's student, Department of Information Technologies of Automated Systems, Faculty of Information Technologies and Management of BSUIR

V.A. Zahariev

Associate Professor of the Department of Control Systems, Faculty of Information Technologies and Management of BSUIR, PhD of Technical Sciences, Associate Professor

Abstract. The review of architectural solutions for distributed streaming data processing systems aimed at speech recognition tasks is presented. The main components of such systems are considered, including the architectural structure of the part, as well as software implementation. Particular attention is paid to using Apache Kafka as a message broker to ensure efficient data transfer. It also describes how to use MLflow to deploy the Whisper speech recognition model, which allows for easy management of the model's lifecycle and metadata. A microservice architecture based on Kubernetes is considered, providing scalability and application management. The results obtained allow us to determine the key directions for the development of distributed streaming data processing systems for efficient and accurate speech recognition in real time.

Keywords: microservice architecture, message broker Apache Kafka, Kubernetes, Whisper, MLflow, speech recognition, distributed systems.

УДК 004; 004.93

ОЦЕНКА КАЧЕСТВА ИЗОБРАЖЕНИЙ ПОСЛЕ ИХ СЖАТИЯ АЛГОРИТМОМ JPEG



Ю.И. Голуб

Старший научный сотрудник
ОИПИ НАН Беларуси, кандидат
технических наук, доцент
6423506 @gmail.com

Ю.И. Голуб

Окончила Белорусский национальный технический университет. Область научных интересов связана с разработкой методов и алгоритмов обработки и анализа цифровых изображений.

Аннотация. Описаны исследования количественных оценок качества, которые были использованы для анализа изображений после их сжатия алгоритмом *JPEG*. Рассматриваются безэталонные оценочные функции качества (no-reference measure). Оценка качества изображений после сжатия влияет на выбор оптимальных настроек различных методов и позволяет улучшать визуальные характеристики сжатого изображения, сохраняя его приемлемое качество. Качество оценивается с точки зрения анализа резкости, контраста и яркости цифровых изображений. Предлагается использовать новый подход к оценке качества изображений, основанный на анализе распределения локальных значений оценочных функций. Показано, что предложенный подход позволяет точнее выполнять оценку качества изображений после их сжатия.

Ключевые слова: цифровое изображение, алгоритм сжатия *JPEG*, количественная оценка качества, распределение Вейбулла.

Введение. Оценка качества цифровых изображений после сжатия алгоритмом *JPEG* является важным этапом в области цифровой обработки изображений. Данные неудовлетворительного качества снижают точность работы алгоритмов обработки и анализа регистрируемых снимков, занимают память, отнимают время у эксперта при их оценке или исследовании. *JPEG* (*The Joint Photographic Experts Group*) – один из наиболее распространённых на сегодня алгоритмов сжатия цифровых изображений [1, 2], который используется при передаче файлов, хранении в базах данных и т.п. Однако при использовании этого алгоритма могут возникать различные искажения, а некоторые детали могут быть утеряны, что сказывается на качестве снимков и их общем восприятии. Поэтому важно выполнять анализ качества изображения после сжатия, чтобы убедиться, что оно соответствует требованиям и ожиданиям пользователей. Сжатие изображений – это поиск компромисса между визуальным качеством и степенью сжатия, или размером выходных данных. Изображения с однородными областями могут сжиматься до меньшего размера и иметь меньше различных артефактов, чем изображения с большим количеством объектов. Для изображений с мелкими деталями или текстом степень сжатия имеет важное значение.

Этап анализа изображений после их сжатия важен для определения уровня потери качества и оценки визуального восприятия изображений в таких областях, как:

– выбор методов для хранения, передачи или шифрования потока данных. Из-за процедуры квантования при сжатии с потерями, структура изображения может быть сильно изменена, особенно при низкой скорости передачи данных, что приводит к очевидному ухудшению качества восприятия, например, к появлению блочности, искажению деталей объектов, размытию краев и т.д.;

– передача больших объемов данных, например, при дистанционном зондировании Земли (ДЗЗ) или дистанционной медицинской диагностике. В этих областях сжатие цифровых изображений представляет серьезную проблему, т.к. необходимо устранить избыточность многоспектральных данных, сократить их объем для оперативной передачи и выполнить сжатие таким образом, чтобы иметь возможность в случае необходимости восстановить исходное изображение.

На рисунке 1 представлены фрагменты исходных изображений и их копий после сжатия алгоритмом *JPEG*. Первое изображение – это спутниковый снимок, полученный при ДЗЗ, второе – снимок сетчатки глаза, используемый для диагностики диабетической ретинопатии.

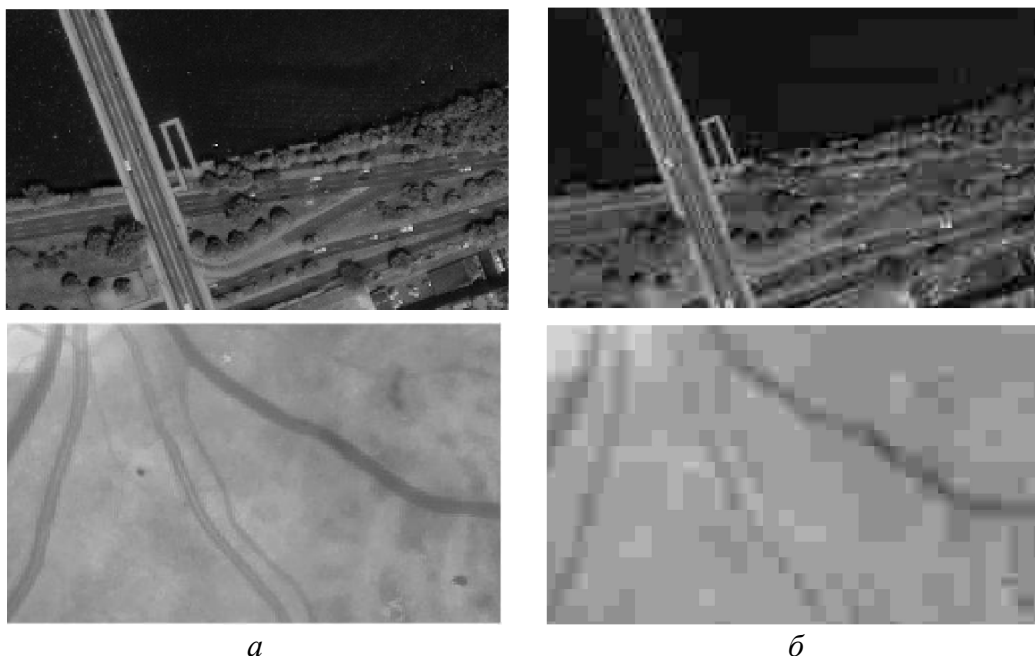


Рисунок 1. Фрагменты изображений: *a* – без сжатия, *б* – высокая степень сжатия алгоритмом *JPEG* (10)

Анализируя изображения рисунка 1, можно заметить, что сжатие оказалось слишком сильным и это привело к потере информации, что может быть крайне критичным в некоторых областях, например, при диагностике заболеваний или ДЗЗ. Оценка качества изображения после сжатия позволяет выбрать оптимальные параметры сжатия для конкретных потребностей и задач, в том числе обеспечить сохранение важных деталей на изображении.

Постановка цели исследования. Для того чтобы оценить качество сжатого снимка, разработаны различные методы и количественные оценки, большинство из которых используют эталон для оценки сходства двух изображений (исходное и сжатое) [3, 4]. Однако не всегда доступно эталонное изображение и необходимо оценить один снимок или их последовательность. Это более сложная задача, при решении которой используются абсолютные (безэталонные или *no-reference*) оценки качества. В реальных прикладных задачах чаще всего используются безэталонные количественные оценки

качества, которые вычисляются локально различными способами. Например, на основе локальных градиентов, статистических характеристик, вейвлет-преобразований и др. Это позволяет оценить качество снимка, учитывая его изменения в различных областях.

Целью написания статьи было исследование количественных безэталонных оценок качества и поиск тех, которые позволяют выбирать изображения хорошего качества и соответствуют визуальному восприятию изображения после его сжатия алгоритмом JPEG.

В настоящее время в исследованиях в области обработки и анализа цифровых изображений широко используются методы машинного обучения (МО). Однако реальные прикладные задачи часто решаются в условиях ограничения объема данных и отсутствия размеченного набора изображений, поэтому методы МО в данном исследовании не рассматривались.

Результаты экспериментального анализа. Для экспериментальных исследований был использован пакет прикладных программ *MATLAB*, в котором реализованы алгоритм сжатия JPEG и стандартные количественные оценки качества изображений, которые можно применять без использования эталона – *BRISQUE*, *PIQE*, *NIQE* [1]. Для указанных оценок, чем меньше значение, тем лучше должно быть качество изображения. На рисунке 2 представлены графики безэталонных количественных оценок, вычисленных для двух изображений, фрагменты которых представлены на рисунке 1. По вертикальной оси графиков – нормализованные значения оценок в диапазоне [0; 1], по горизонтальной оси – степень сжатия алгоритмом JPEG, где 10 – это максимальная степень сжатия (результат сжатия на рисунке 1б), 100 – это минимальная степень сжатия, а значением 110 обозначено исходное изображение без сжатия.

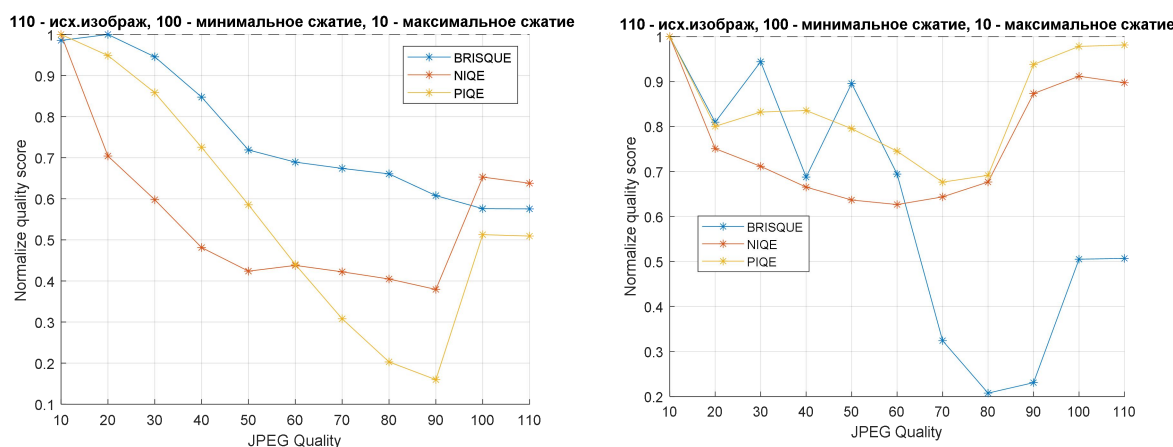


Рисунок 2. Количественные оценки, используемые в *MATLAB* для анализа качества изображений без наличия эталона

Проводя анализ полученных экспериментальных данных, можно отметить, что количественные оценки качества *BRISQUE*, *PIQE*, *NIQE* для определенных типов изображений не соответствуют визуальному восприятию человека. Такой пример представлен на рисунке 2, где минимальное значение оценок не соответствует изображению с лучшим качеством, т.е. исходному изображению. Аналогичные результаты, где количественные оценки *BRISQUE*, *PIQE*, *NIQE* не соответствуют визуальному анализу, были получены для изображений из общедоступной базы *TID2013* [5] после их сжатия алгоритмом *JPEG*. На рисунке 3 представлено одиннадцать анализируемых изображений (одно исходное *I01_jpeginput_image.jpg* из базы *TID2013* и десять его копий с разной степенью сжатия). Т.к. при печати сложно заметить разницу изображений рисунка 3, на рисунке 4 представлены увеличенные фрагменты шести изображений. На рисунке 5 показаны графики оценочных функций: три стандартные из

среды *MATLAB* и функция *NJQA* [6], используемая для оценки качества изображений после сжатия.

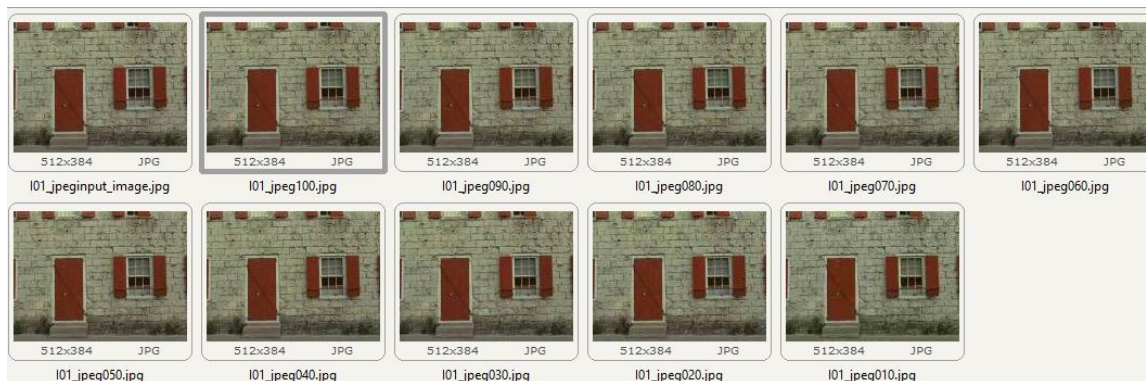


Рисунок 3. Последовательность из 11 изображений, где первое изображение – это исходное изображение из базы TID2013 (без сжатия), остальные – после сжатия алгоритмом *JPEG* (слева направо степень сжатия увеличивается от 100 до 10)

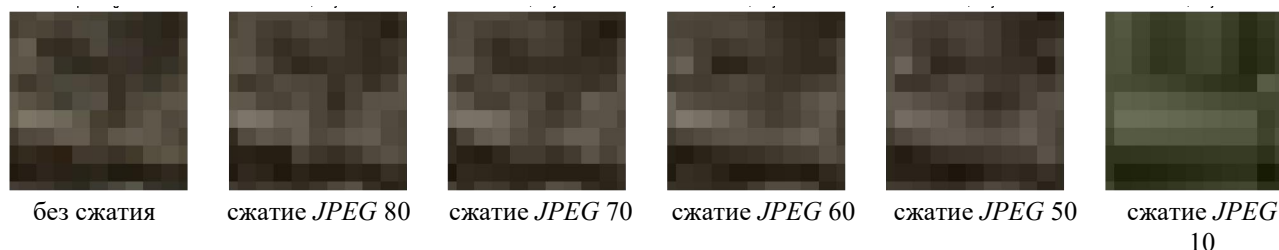


Рисунок 4. Увеличенные фрагменты изображений, представленных на рисунке 3 (слева направо степень сжатия увеличивается, последнее изображение с максимальным сжатием)

На графиках рисунка 5 можно заметить, что количественные оценки изображений не соответствуют визуальным, т.к. исходный снимок без сжатия должен иметь минимальные значения.

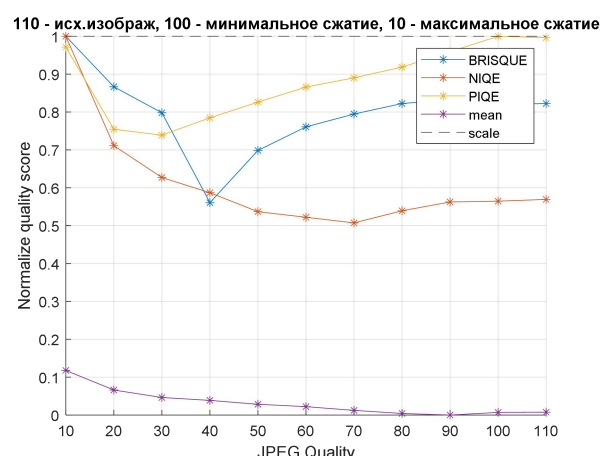


Рисунок 5. Количественные оценки качества изображений, представленных на рисунке 3

Экспериментальный анализ показал, что чем меньше разрешение изображения в пикселях, тем меньше его можно сжимать без потери визуального качества. При сжатии изображения меньшего разрешения, искажения проявляются в большей степени, но и размер изображения сокращается больше (рисунок 6). Таким образом, при выборе между

сжатием изображения меньшего разрешения и сжатием изображения большего разрешения, необходимо учитывать, как визуальное качество, так и размер файла. В зависимости от задачи и требований пользователей, можно найти оптимальный баланс между масштабированием изображения и сохранением деталей, чтобы достичь желаемого результата.

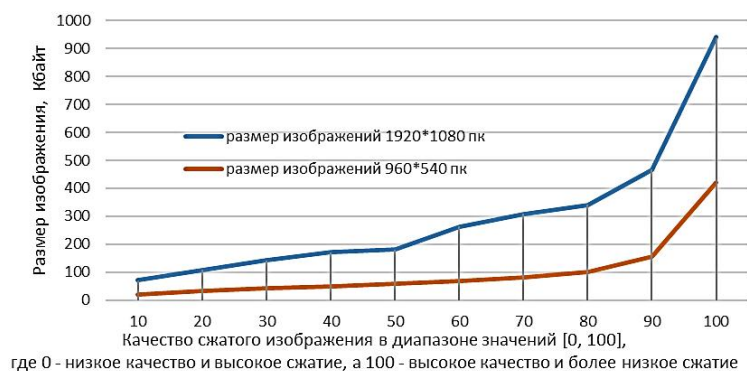


Рисунок 6. Изменение размера изображения в зависимости от качества файла, сжатого в формате *JPEG*

Было проанализировано тридцать безэталонных оценочных функций качества [1], которые используют усреднение локальных значений и упоминаются в отечественной, а также в зарубежной литературе. Оценки вычислялись как в частотной области, так и в пространственной. Размер локальной области мог быть фиксированным (например, 15x15 пикселей) или зависел от размера анализируемого изображения. Обычно в исследованиях используют предположение, что локальные оценки качества, вычисленные на основе одного изображения, нормально распределены, и в роли глобальной оценки применяют параметры нормального распределения – среднее арифметическое (математическое ожидание) и среднеквадратичное локальных оценок. Как показывают эксперименты распределение локальных значений оценочных функций чаще всего не соответствует нормальному распределению. Пример изображения после сжатия *JPEG* и его гистограмма локальных оценок функции *NJQA* представлены на рисунке 7. Анализ гистограмм локальных оценок (на нескольких тысячах изображений) показал, что они могут иметь несимметричную и мультимодальную форму распределения.

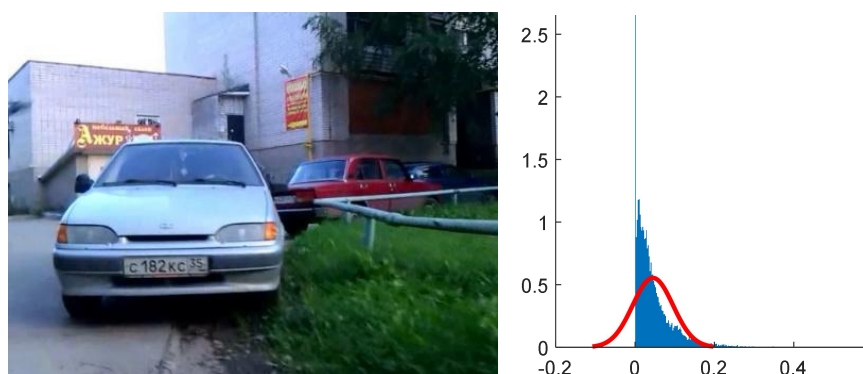


Рисунок 7. Изображение после сжатия и распределение его локальных значений оценочной функции качества *NJQA*. Визуальный анализ показывает, что форма гистограммы не соответствует нормальному распределению данных (красная линия на графике)

Для выбора более подходящего варианта аппроксимации множества локальных значений оценочных функций было исследовано 16 распределений данных: равномерное, нормальное, логарифмическое нормальное (логнормальное), экспоненциальное, гамма, Вейбулла, логистическое, логарифмическое логистическое, обобщенное распределение экстремальных значений, Бирнбаума-Саундерса, экстремальных значений, обобщенное Парето, обратное Гаусса, Накагами, Рэлея, Стьюдента. Анализ экспериментальных данных показывает, что распределение Вейбулла, имея более гибкую форму, точнее описывает множество локальных оценок, вычисленных для изображения. Предлагается вместо усреднения использовать параметры распределения Вейбулла – масштаб (*scale*) и форма (*form*). Для демонстрации предлагаемого подхода на рисунке 8 представлены графики для трех вариантов обобщения локальных значений оценочной функций качества *GLLV* [1], вычисленные для спутникового снимка (фрагмент его представлен на рисунке 1а) и его копий, полученных сжатием разной степени. Вертикальная и горизонтальная оси графиков соответствуют описанию осей рисунка 2. Анализ графиков показывает, что параметры распределения Вейбулла предпочтительнее использовать для обобщения локальных значений оценочных функций, а параметр *mean* (среднее) не соответствует визуальной оценке человека.

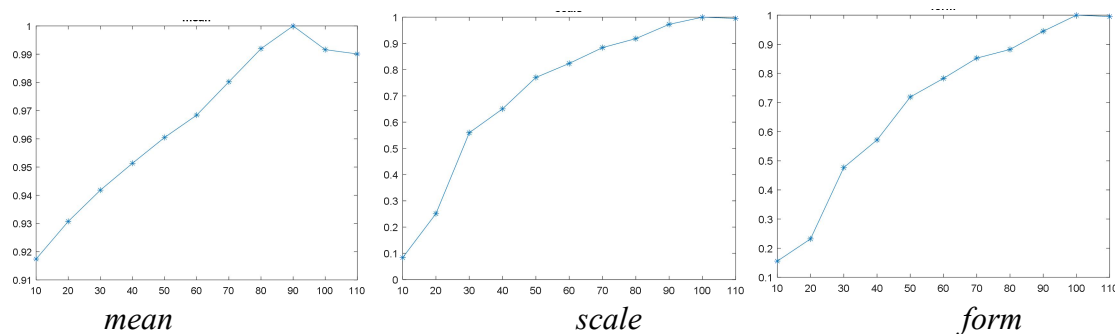


Рисунок 8. Графики для трех вариантов обобщения локальных значений количественной оценки *GLLV* (по горизонтальной оси – степень сжатия *JPEG*, по вертикальной оси – нормализованные значения оценок)

Заключение. Применение оперативной и точной автоматической оценки качества позволяет исключать из базы данных изображения неудовлетворительного качества, тем самым сокращая объем хранилища и экономя время специалистов, занимающихся последующим анализом данных. Это особенно актуально в областях, где требуется обработка большого количества цифровых изображений, таких как медицинская диагностика, научные исследования и промышленность.

Эксперименты показали, что использование распределения Вейбулла для моделирования количественных оценок изображений дает более точные результаты в оценке качества, чем нормальное распределение.

Работа частично выполнена в рамках проекта № Ф22СРБГ-001.

Список литературы

- [1] Голуб Ю. И. Оценка качества цифровых изображений / Ю.И. Голуб, В.В. Старовойтов. – Минск : ОИПИ НАН Беларуси, 2023. – 252 с.
- [2] ITU-T T.81 (1992) | ISO/IEC 10918-1:1994, Information technology – Digital compression and coding of continuous-tone still images – Requirements and guidelines. – 182p.
- [3] Zhang W., Ma K., Zhai G., Yang X. Uncertainty-Aware Blind Image Quality Assessment in the Laboratory and Wild // IEEE Transactions on Image Processing. – 2021. – Vol.30. – P.3474–3486.
- [4] Zhang X., Kwong S., Kuo C.-C.J. Data-Driven Transform-Based Compressed Image Quality Assessment // IEEE Transactions on Circuits and Systems for Video Technology. – 2021. – 31(9). – P.3352–3365.

[5] Color image database TID2013: Peculiarities and preliminary results / N. Ponomarenko [et al.] // Signal Processing: Image Communication. – 2015. – Vol.30. – P.57–77.

[6] No-reference quality assessment of JPEG images via a quality relevance map // Signal Processing Letters. – 2014. – Vol.21. – №.2. – P.155–158.

IMAGE QUALITY ASSESSMENT AFTER JPEG COMPRESSION

Y.I. Golub

*Ph.D., Associate Professor, Senior
Researcher UIIP NAS of the
Republic of Belarus*

Abstract. The article describes the research of quality assessment methods, which were used to analyze images after their JPEG compression. No-reference measures are considered. Quality assessment of images after compression impacts the choice of optimal settings of various methods and allows to improve the visual characteristics of the compressed image while maintaining its acceptable quality. The quality is evaluated in terms of analyzing the sharpness, contrast and brightness of digital images. It is proposed to use a new approach to image quality estimation based on the analysis of distribution of local values of estimation functions. It is shown that the proposed approach allows more accurate estimation of image quality after image compression.

Keywords: digital image, JPEG compression algorithm, quality assessment, Weibull distribution.

УДК 004.021:004.75.

ОПТИМИЗАЦИЯ ЧЕЛОВЕКО-МАШИННЫХ ВЗАИМОДЕЙСТВИЙ ЧЕРЕЗ ИННОВАЦИОННЫЕ СИСТЕМЫ СТИМУЛИРОВАНИЯ ТРУДОВОЙ ДЕЯТЕЛЬНОСТИ



Ф.В. Усенко
инженер-программист
кафедры инженерной
психологии и эргономики
БГУИР, магистр
f.usenko@bsuir.by



М.С. Ильясова
ассистент кафедры инженерной
психологии и эргономики БГУИР,
магистр
m.iliasova@bsuir.by

М.С. Ильясова

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов находится на пересечении информационных технологий и промышленной психологии.

Ф.В. Усенко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Областью научных интересов является моделирование пользователей информационных систем и разработкой интерфейсов.

Аннотация. В статье исследуется влияние систем стимулирования трудовой деятельности на эффективность взаимодействия операторов с человеко-машинными системами. Акцент сделан на адаптации таких систем под индивидуальные особенности пользователей и интеграцию современных технологий, таких как искусственный интеллект и машинное обучение, для повышения мотивации и производительности. Обсуждаются проблемы и вызовы, связанные с внедрением этих систем, и предлагаются пути их решения.

Ключевые слова: Человеко-машинные системы, стимулирование трудовой деятельности, искусственный интеллект, машинное обучение, мотивация, производительность, индивидуализация, эффективность, внедрение технологий.

Введение. В современном мире человеко-машинные системы играют ключевую роль в различных областях человеческой деятельности, начиная от промышленного производства и заканчивая высокотехнологичными исследованиями. Важность эффективного взаимодействия между оператором и машиной не может быть переоценена, поскольку от этого напрямую зависят продуктивность труда, безопасность процессов и качество конечного продукта. В этом контексте, система стимулирования трудовой деятельности операторов приобретает особую актуальность, поскольку она направлена на максимизацию эффективности работы и повышение уровня удовлетворенности работой [1-2].

Важным аспектом эффективности человеко-машинных систем является глубокое понимание взаимодействия оператора и машины. Это взаимодействие охватывает не только технические, но и психологические, эргономические и социальные аспекты.

Трудовая деятельность в контексте ЧМС представляет собой совокупность умственных и физических усилий оператора, направленных на управление машиной и обработку

получаемой от нее информации. Стимулирование трудовой деятельности включает в себя меры, направленные на повышение мотивации и эффективности работы оператора, среди которых можно выделить материальное стимулирование, повышение интереса к работе, развитие профессиональных навыков и создание комфортных условий труда.

Обзор существующих подходов к стимулированию. Существует множество подходов к стимулированию трудовой деятельности, включая теории мотивации, организационную психологию и управление человеческими ресурсами. Важными являются работы А. Маслоу, Ф. Герцберга, Д. Макклелланда, которые исследуют различные аспекты мотивации и ее влияние на трудовую активность [3].

В контексте ЧМС особое внимание уделяется таким аспектам, как адаптация интерфейсов к потребностям и предпочтениям операторов, минимизация умственной нагрузки и предотвращение профессионального выгорания. Эргономика и дизайн рабочего места играют ключевую роль в создании оптимальных условий для работы оператора.

Существующие подходы к стимулированию трудовой деятельности операторов человеко-машинных систем можно классифицировать по нескольким основным направлениям:

1 Мотивационные теории.

Теория потребностей Абрахама Маслоу предполагает, что мотивация человека проистекает из стремления удовлетворить различные уровни потребностей: от физиологических до самореализации.

Теория двух факторов Фредерика Герцберга разделяет факторы на «гигиенические» (условия труда, заработная плата и т.д.), которые не мотивируют, но могут вызвать недовольство, и «мотиваторы» (достижение, признание, работа сама по себе), которые способствуют высокой мотивации и удовлетворенности.

Теория ожидания Виктора Врума утверждает, что мотивация зависит от ожидаемой вероятности достижения желаемого результата и ценности этого результата для индивида.

2 Организационные подходы.

Системы вознаграждений и компенсаций предполагают использование материальных и нематериальных стимулов для повышения мотивации и производительности работников.

Участие в принятии решений и делегирование полномочий могут увеличить вовлеченность и мотивацию сотрудников за счет предоставления им большей автономии и ответственности [4].

Карьерное развитие и обучение предлагают долгосрочную перспективу для роста и развития сотрудников, что повышает их лояльность и мотивацию.

3 Психологические и эргономические подходы

Оптимизация условий труда и рабочего пространства направлена на снижение физической и умственной усталости, предотвращение профессионального выгорания и повышение общей удовлетворенности работой.

Обратная связь и признание важны для поддержания мотивации, поскольку операторы человеко-машинных систем часто работают в условиях, где результаты их труда не всегда немедленно видны [5].

Игровые элементы и геймификация в процессе работы могут увеличить вовлеченность и мотивацию за счет введения элементов конкуренции, достижений, наград и прогресса.

Эти подходы могут быть адаптированы и интегрированы в систему стимулирования трудовой деятельности операторов человеко-машинных систем, учитывая специфику работы, индивидуальные особенности сотрудников и цели организации.

Междисциплинарные аспекты взаимодействия человека и машины. К междисциплинарным аспектам, влияющим на взаимодействие человека и машины, относят:

Этические соображения: Обсуждение этических вопросов, связанных с автоматизацией и ролью человека в человеко-машинных системах. Включение проблематики

ответственности за принятие решений и возможных социальных последствий автоматизации [6].

Культурные и социальные аспекты: Анализ того, как культурные и социальные различия влияют на взаимодействие человека и машины. Исследование влияния культурных предпочтений и традиций на проектирование и восприятие человеко-машинных интерфейсов.

Тенденции и новые направления в развитии человеко-машинных систем. С развитием технологий интерфейсов, таких как голосовые ассистенты, жестовое управление и аугментированная реальность, взаимодействие между человеком и машиной становится более интуитивным и естественным. Эти технологии позволяют создавать более адаптивные и понятные системы управления, которые могут предугадывать намерения пользователя и облегчать его работу с машиной.

Искусственный интеллект и машинное обучение.

Применение ИИ и машинного обучения в человеко-машинных системах позволяет системам адаптироваться к поведению и предпочтениям пользователя, оптимизируя процессы и делая взаимодействие более эффективным. ИИ может анализировать большие объемы данных для предсказания необходимых действий и предоставления рекомендаций, тем самым улучшая качество работы и снижая вероятность ошибок.

Робототехника и автоматизация.

Развитие робототехники приводит к созданию более сложных и автономных систем, способных выполнять задачи с минимальным вмешательством человека. Это особенно актуально в опасных или труднодоступных средах, где роботы могут заменить человека, снижая риски и увеличивая эффективность выполнения задач.

Киберфизические системы и Интернет вещей (IoT).

Интеграция киберфизических систем и IoT приводит к созданию умных сред, где человеко-машинное взаимодействие расширяется за счет обмена данными между различными устройствами и системами. Это позволяет создавать комплексные экосистемы, в которых машины могут синхронизироваться и координировать свою работу, предоставляя пользователю более глубокий контроль и лучшее понимание процессов[7].

Умные и адаптивные интерфейсы.

Разработка умных интерфейсов, способных адаптироваться к индивидуальным особенностям и предпочтениям пользователя, является одним из важных направлений в области человеко-машинных систем. Такие интерфейсы могут изменять свою структуру и представление информации в зависимости от контекста использования, опыта пользователя и специфических задач, тем самым улучшая удобство использования и эффективность взаимодействия.

Влияние технологических инноваций на развитие систем стимулирования.

Технологические инновации играют ключевую роль в трансформации систем стимулирования трудовой деятельности, особенно в контексте человеко-машинных систем. Новые технологии не только расширяют возможности для мотивации и повышения производительности, но и в корне меняют подходы к организации трудового процесса.

Блокчейн и децентрализованные системы.

Блокчейн-технологии предлагают новые методы для создания прозрачных, справедливых и автоматизированных систем вознаграждения и мотивации. Использование смарт-контрактов позволяет автоматически выплачивать вознаграждения за достижение определенных KPI или выполнение специфических задач, устраняя необходимость в посредниках и снижая вероятность ошибок и злоупотреблений.

Искусственный интеллект и аналитика данных.

ИИ и аналитика данных предоставляют возможности для глубокого анализа производительности, предпочтений и поведения операторов, позволяя создавать персонализированные и адаптивные системы стимулирования. Алгоритмы машинного

обучения могут выявлять закономерности и предоставлять рекомендации по оптимизации рабочих процессов и мотивационных стратегий. рамках симулятора грузового автомобиля каждое взаимодействие пользователя с виртуальным автомобильным миром зависит от разработанных сценариев, которые стремятся обеспечить уникальный и иммерсивный опыт. Система скриптов взаимодействует с каждым аспектом виртуальной среды, гарантируя высокую степень реализма в интерактивных процессах [8].

Нейротехнологии и интерфейсы мозг-компьютер.

Развитие нейротехнологий и интерфейсов мозг-компьютер открывает новые перспективы для понимания и стимулирования трудовой деятельности на уровне нейронных процессов. Эти технологии могут помочь в идентификации уровней усталости, стресса, внимания и мотивации операторов в реальном времени, позволяя корректировать нагрузку и стимулирование для оптимизации работы.

Виртуальная и дополненная реальность.

Технологии виртуальной и аугментированной реальности предлагают уникальные возможности для обучения и развития навыков операторов в условиях, максимально приближенных к реальным. Использование этих технологий в системах стимулирования может сделать процесс обучения более интерактивным и мотивирующим, а также позволить операторам отрабатывать сложные сценарии без риска для здоровья и безопасности.

Мобильные технологии и приложения.

Мобильные устройства и приложения позволяют интегрировать системы стимулирования непосредственно в повседневную жизнь операторов, предоставляя инструменты для самомониторинга, обучения и получения обратной связи в любое время и в любом месте. Это делает процесс мотивации более гибким и доступным.

Таким образом, технологические инновации не только улучшают существующие системы стимулирования, но и создают новые подходы к взаимодействию человека и машины, способствуя повышению эффективности и удовлетворенности трудовой деятельности.

Проблемы и вызовы при внедрении систем стимулирования.

Сопротивление изменениям.

Сотрудники и менеджмент могут проявлять сопротивление новым системам стимулирования из-за привычки к существующим методам работы или боязни потери контроля и автономии.

Интеграция с существующими системами.

Трудности интеграции новой системы стимулирования с уже используемыми в организации информационными и управленческими системами могут привести к техническим проблемам и сбоям в работе.

Недостаток персонализации.

Отсутствие гибкости и персонализации в системе стимулирования может привести к недовольству среди сотрудников, поскольку различные люди мотивируются по-разному.

Оценка эффективности.

Сложность объективной оценки влияния системы стимулирования на производительность и мотивацию сотрудников, особенно в краткосрочной перспективе.

Этические и правовые вопросы.

Вопросы конфиденциальности и этики при сборе и анализе данных о сотрудниках, а также соответствие системы стимулирования трудовому законодательству и стандартам.

Бюджетные ограничения.

Высокая стоимость разработки и внедрения систем стимулирования может стать препятствием для малых и средних предприятий. виртуальной среды.

Адаптация к изменяющимся условиям.

Необходимость постоянной адаптации системы стимулирования к изменяющимся внутренним и внешним условиям бизнеса и трудового коллектива.

Решение этих проблем требует комплексного подхода, включая активное вовлечение сотрудников в процесс разработки и внедрения системы, гибкость системы и возможность её адаптации под конкретную организационную культуру и специфику работы, а также тщательное планирование и контроль за процессом внедрения на всех этапах.

Заключение. В ходе исследования было подчеркнута ключевое значение человеко-машинных систем и систем стимулирования трудовой деятельности для повышения эффективности и удовлетворенности работы операторов. Современные технологические инновации, включая ИИ, блокчейн и нейротехнологии, открывают новые горизонты для развития этих систем, делая их более адаптивными и персонализированными. Однако внедрение таких систем сталкивается с вызовами, включая сопротивление изменениям и этические вопросы, требующие комплексного и включающего подхода для их преодоления и достижения гармоничного взаимодействия человека и машины.

Список литературы

- [1] Макаров А.В. Конончук Е.Н. Эргономика и психология взаимодействия с компьютером: учебник. – М.: Юрайт 2021.
- [2] Lazar J., Feng J. H., & Hochheiser H. Research Methods in Human-Computer Interaction. Morgan Kaufmann.
- [3] Новикова О.П., Попов А.С. Психологические аспекты создания и использования виртуальных интерфейсов. // Психология виртуальной реальности и киберпсихология. – 2022. – Т. 3 № 1.
- [4] Johnson J. The psychology of user interface design: A brief guide. Psychology Today. [Online]. Available: Psychology Today.
- [5] Dix A., Finlay J., Abowd G., & Beale R. Human-computer interaction. Pearson.
- [6] Nielsen J. Usability 101: Introduction to usability. Nielsen Norman Group. [Online]. Available: Nielsen Norman Group.
- [7] Norman D. A. The design of everyday things: Revised and expanded edition. Basic Books.
- [8] Lazar J., Feng J. H., & Hochheiser H. (2017). Research Methods in Human-Computer Interaction. Morgan Kaufmann.

Авторский вклад

Авторы внесли равноценный вклад.

OPTIMIZATION OF HUMAN-MACHINE INTERACTIONS THROUGH INNOVATIVE WORK INCENTIVE SYSTEMS

P.V. Usenko
Software engineer, Department
of Engineering Psychology and
Ergonomics BSUIR, M.Sc.
f.usenko@bsuir.by

M.S. Ilyasova
Assistant of the Department of
Engineering Psychology and
Ergonomics of BSUIR, M.Sc.

Abstract. The article examines the impact of labor incentive systems on the effectiveness of operator interaction with human-machine systems. The emphasis is on adapting such systems to the individual characteristics of users and integrating modern technologies such as artificial intelligence and machine learning to increase motivation and productivity. The problems and challenges associated with the implementation of these systems are discussed and ways to solve them are proposed.

Keywords: Human-machine systems, stimulation of labor activity, artificial intelligence, machine learning, motivation, productivity, individualization, efficiency, technology implementation.

УДК 004.021:004.75

ПРИЛОЖЕНИЕ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ «ТУР ПО МУЗЕЯМ»



М.Н. Мельников
Учащийся Национального
детского технопарка,
учащийся гимназии № 1
г. Солигорска



И.Д. Полегин
Учащийся Национального
детского технопарка,
учащийся гимназии № 10
г. Гомеля



Ф.В. Усенко
инженер-программист
кафедры инженерной
психологии и эргономики
БГУИР, магистр
f.usenko@bsuir.by



М.С. Ильясова
ассистент кафедры
инженерной психологии и
эргономики БГУИР,
магистр
m.iliasova@bsuir.by



Л.Р. Коркин
Заведующий лабораторией
виртуальной и дополненной
реальности Национального
детского технопарка,
магистр технических наук



А.М. Прудник
доцент кафедры инженерной
психологии и эргономики
БГУИР, кандидат
технических наук, доцент
aleksander.prudnik@bsuir.by

М.Н. Мельников

Обучается в гимназии № 1 г.Солигорска. Область научных интересов связана с технологиями виртуальной и дополненной реальности.

М.С. Ильясова

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов находится на пересечении информационных технологий и промышленной психологии.

Ф.В. Усенко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Областью научных интересов является моделирование пользователей информационных систем и разработкой интерфейсов.

Л.Р. Коркин

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с системами распознавания изображений.

А.М. Прудник

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с взаимодействием человека с компьютером, интерфейсами информационных систем, front-end web development.

Аннотация. Исследуется влияние информационных технологий на взаимодействие с культурными артефактами, с акцентом на виртуальные музеи как новый этап цифровой трансформации. Описываются различные типы взаимодействия между посетителями и виртуальными объектами, а также их вклад в улучшение впечатлений посетителей и понимания культурного наследия. Рассматриваются инструменты разработки VR-приложений, включая *Unity*, *C#* и *Blender*, и их роль в создании интерактивного приложения. Описывается процесс реализации VR-приложения, включающего выбор артефакта для реставрации, знакомство с инструментами и материалами, применение различных методов реставрации и создание бесконечного цикла обучения. Подчеркивается, что приложение способствует вовлечению и мотивации пользователей, обучению, самовыражению и творчеству, а также улучшению навыков и знаний в области реставрации. Использование графических ресурсов движка *Unity* обеспечивает высокий уровень визуальной реалистичности и эффект полного погружения. Успешная реализация проекта подтверждает эффективность VR-приложения для реставрации культурных артефактов в качестве инструмента обучения и развития навыков в области реставрации, а также способствует популяризации сохранения культурного наследия.

Ключевые слова: виртуальная реальность, *unity*, *c#*, объектно-ориентированное программирование, 3d-моделирование, кроссплатформенная разработка, интерактивная среда, рендеринг в реальном времени, реставрация культурных артефактов, образование в области культурного наследия, визуализация исторического контекста.

Введение. Возникновение VR открыло новые возможности для сохранения и восстановления культурного наследия. Эта технология предлагает интерактивную платформу для обучения и популяризации охраны культурного наследия. Применение VR в реставрации культурных артефактов является развивающимся направлением исследований, которое предоставляет инновационный подход к пониманию, анализу и восстановлению культурных ценностей. Создание виртуальной копии артефакта позволяет реставраторам исследовать и взаимодействовать с ним в новых способах. Такой подход позволяет более детально изучить артефакт и повысить эффективность его реставрации. Кроме того, VR может служить полезным образовательным инструментом для реставраторов, предоставляя им практический опыт в безопасной среде. С помощью VR, реставраторы могут получить более глубокое понимание исторического контекста артефакта и его физических свойств, что поможет им выбрать наилучшие методы реставрации. Кроме того, виртуальная реальность является важным инструментом в популяризации охраны культурного наследия, предлагая захватывающий опыт взаимодействия с культурными ценностями. Это поможет повысить осведомленность общественности о важности сохранения культурного наследия, позволяя людям виртуально исследовать исторические места и артефакты.

Анализ взаимодействия посетителей виртуальных музеев с культурными объектами и постановка задачи. Развитие информационных технологий изменило то, как мы взаимодействуем с культурными артефактами и музеями. Виртуальные музеи и артефакты стали новым этапом в цифровой трансформации, предлагая посетителям захватывающий и интерактивный опыт, выходящий за пределы ограничений физического пространства. Необходимо классифицировать различные типы взаимодействия, которые происходят между посетителями и этими виртуальными объектами, исследуя, как эти взаимодействия улучшают впечатления посетителей, облегчают обучение и способствуют более глубокому пониманию культурного наследия.

Взаимодействие посетителей с виртуальными музеями и артефактами можно разделить на несколько видов. Взаимодействие с платформой или устройством, что предполагает взаимодействие с платформой или виртуальным устройством, в основном посредством жестов, прикосновений, звуков и других действий посетителя [1]. Взаимодействие на основе датчиков, что предполагает использование датчиков для сбора данных, например датчики движения для отслеживания движений пользователя [2]. Взаимодействие на основе устройств, что предполагает использование устройств ввода, таких как мышь, клавиатура или сенсорный экран, для взаимодействия с виртуальной средой [3]. Материальное взаимодействие, что подразумевает взаимодействия

с физическими объектами, дополненными или соединенными цифровым способом [2]. Совместное взаимодействие., что подразумевает взаимодействие нескольких пользователей с виртуальной средой или друг с другом [3]. Мультимодальное взаимодействие, что предполагает одновременное использование нескольких режимов ввода, например голоса и жестов [4]. Гибридное взаимодействие, которое сочетает в себе несколько вышеперечисленных методов, обеспечивая наиболее захватывающий опыт [4]. Все эти методы взаимодействия используются технологиями иммерсивной реальности для обеспечения взаимодействия с виртуальной средой. Они существенно меняют опыт посетителей музея и сам процесс осмысления музейных коллекций.

Спрос на виртуальные музеи и галереи значительно увеличился, особенно после глобального кризиса в области здравоохранения, вызванного *COVID-19* [5]. Поскольку миллионы людей оказались заперты в закрытых помещениях из-за карантина, спрос на виртуальное посещение музеев резко возрос. Интеграция цифровых платформ в музеи и галереи ускорила, что изменило способы взаимодействия посетителей с контентом культурного наследия.

Конкретные данные о доле виртуальных музеев среди приложений виртуальной реальности во всём мире отсутствуют. Доступны только сведения по отдельным странам. Так, например, в 2020 году только 11% музеев, галерей и аналогичных культурных учреждений Италии предлагали своим посетителям инструменты виртуальной реальности, однако при этом 50% учреждений планировали внедрить эту технологию в будущем [6]. Что касается мирового рынка виртуальной реальности, то он, по многим прогнозам, существенно вырастет. Так, суммарный объем рынка виртуальной реальности оценивался в \$19,44 млрд в 2022 году и, прогнозируется его рост до \$165,91 млрд к 2030 году [7]. Учитывая растущий спрос на виртуальные музеи, можно сделать вывод, что они составляют значительную часть этого рынка.

Было обнаружено, что виртуальные музеи повышают интерес к реальным объектам. Они предоставляют расширенные возможности посещения музея, что повышает вовлеченность и привлекательность посетителей [8]. Более того, они повлияли на многих людей, которые ранее были против посещения этих мест, и добавили их в свой список желаний. Включение в эти музеи дополнительной истории, контекста и техник к шедеврам и менее известным произведениям искусства повысило интерес людей и побудило их однажды посетить эти места лично [9].

Виртуальные музеи играют все более важную роль в сфере приложений виртуальной реальности. Они удовлетворяют растущий спрос на захватывающий и интерактивный

Выбор инструментов. Выбор оптимальных инструментов разработки для создания *VR*-приложения, направленного на восстановление культурных артефактов, включает *Unity* [10], язык *C#* [11] и *Blender*. *Unity* представляет собой мощную платформу, поддерживающую различные *VR*-устройства, имеющую передовые технологии рендеринга и анимации. Его удобный механизм скриптов позволяет адаптировать приложение под потребности реставраторов, а возможность масштабирования обеспечивает стабильность проекта при его расширении. Язык *C#* был выбран для написания скриптов благодаря его мощности, простоте и объектно-ориентированности. Он обеспечивает гладкую интеграцию с *Unity* и позволяет напрямую взаимодействовать с *API Unity*. Строгая типизация *C#* снижает вероятность ошибок, делая код более надежным. *Blender*, который был выбран для создания *3D*-объектов, предлагает целый ряд преимуществ. Он является комплексным инструментом с открытым исходным кодом, который поддерживает полный цикл разработки. Файлы *Blender* легко импортировать в *Unity*, что упрощает рабочий процесс. Все эти факторы делают *Unity*, *C#* и *Blender* наиболее целесообразным выбором для данного проекта.

Реализация приложения. Реализация VR-приложения состояла из нескольких ключевых этапов. На начальном этапе пользователь должен выбрать артефакт для реставрации с помощью интерактивного интерфейса в виртуальном офисе (рисунок 1). Затем пользователь переносится в виртуальную мастерскую, где знакомится с инструментами и материалами, необходимыми для реставрации. На протяжении всего процесса реставрации пользователь последовательно применяет различные инструменты и методы, что требует от него навыков и знаний, связанных с реставрацией произведений искусства. После того как рама отделена от подрамника (рисунок 2), пользователь может детально изучить артефакт, определить места, требующие особого внимания в процессе реставрации (рисунок 3), и спланировать дальнейшие действия. Завершив реставрацию одного артефакта, пользователь может выбрать другой, создавая бесконечный цикл обучения и применения новых навыков.

Игровая механика VR-приложения способствует вовлечению и мотивации пользователей благодаря обучению в процессе взаимодействия, цикличности игрового процесса, самовыражению и творчеству, а также обучению и достижениям. Пользователи учатся и развиваются благодаря активному взаимодействию с виртуальным миром, используя инструменты и методы восстановления. Бесконечный цикл взаимодействия поддерживает долгосрочную мотивацию и вовлеченность. Приложение позволяет пользователям выразить свою креативность и уникальность в процессе реставрации. Использование приложения приводит к значительному улучшению навыков и знаний пользователей в области художественной реставрации. Пользователи приобретают опыт использования различных инструментов и методов реставрации, что способствует повышению их компетенции в этой области. Таким образом, приложение для реставрации культурных артефактов является эффективным инструментом для обучения и развития навыков в области реставрации. В симуляторе используются графические ресурсы движка *Unity*, что обеспечивает высокий уровень визуальной реалистичности. Это способствует созданию эффекта полного погружения при использовании устройств виртуальной реальности.



Рисунок 1. Рабочее место реставратора



Рисунок 2. Отсоединение рамки



Рисунок 3. Снятие загрязнений

Разработка кода. В этом разделе подробно описаны ряд различных скриптов, используемых в проекте. Эти скрипты в целом делятся на два типа: те, которые облегчают прямое взаимодействие между пользователем и объектами/средой, и вспомогательные скрипты, которые поддерживают это взаимодействие.

Например, был написан скрипт `ToolActionHandScript` для взаимодействия рук крепления подрамника к раме картины (рисунок 4). Этот скрипт является классом в *Unity*, который управляет взаимодействием рук с объектами в виртуальной реальности. В классе были определены несколько полей. `tagToInteract` – это тег объектов, с которыми можно взаимодействовать. `isGrabbing` и `isInteracting` – это булевы переменные, которые отслеживают, происходит ли в данный момент захват или взаимодействие. `Hand`, `HandHand`, `HandGripButton` и `HandInteractButton` – это объекты, связанные с библиотекой *SteamVR* и используются для отслеживания ввода пользователя через контроллеры *VR*.

```
main.cs
1 using System.Collections;
2 using System.Collections.Generic;
3 using Cinemachine;
4 using UnityEngine;
5 using Valve.VR;
6 using Valve.VR.InteractionSystem;
7
8 public class ToolActionHandScript : MonoBehaviour
9 {
10     [SerializeField] public string tagToInteract;
11
12     public bool isGrabbing;
13     public bool isInteracting;
14     public SteamVR_Input_Sources Hand;
15     public Hand HandHand;
16     public SteamVR_Action_Boolean HandGripButton;
17     public SteamVR_Action_Boolean HandInteractButton;
18
19     private void Update()
20     {
21         if (HandGripButton.GetState(Hand))
22         {
23             isGrabbing = true;
24         }
25         else
26         {
27             isGrabbing = false;
28         }
29
30         if (HandInteractButton.GetState(Hand))
31         {
32             isInteracting = true;
33         }
34         else
35         {
36             isInteracting = false;
37         }
38     }
39 }
```

Рисунок. 4. Скрипт взаимодействия рук креплениями подрамника к раме картины

```
40
41 private void OnTriggerStay(Collider other)
42 {
43     if (other.tag == tagToInteract)
44     {
45         if (HandInteractButton.GetStateDown(Hand))
46         {
47             if (other.gameObject.TryGetComponent<LeafHolderScript>(out LeafHolderScript
48 leafHolderScript))
49             {
50                 leafHolderScript.Interaction();
51             }
52         }
53     }
54 }
```

Продолжение рисунка 4

Метод Update вызывается каждый кадр и проверяет состояние кнопок HandGripButton и HandInteractButton. Если кнопка HandGripButton нажата, isGrabbing устанавливается в true, в противном случае – в false. Аналогично, если кнопка HandInteractButton нажата, isInteracting устанавливается в true, в противном случае – в false.

Метод OnTriggerStay вызывается каждый кадр, пока другой коллайдер находится внутри триггера. Если тег другого коллайдера соответствует tagToInteract и кнопка HandInteractButton нажата, то вызывается метод Interaction у компонента LeafHolderScript этого объекта, если он существует.

То есть, этот скрипт позволяет пользователю взаимодействовать с определенными объектами в виртуальной реальности, используя контроллеры VR. Когда пользователь нажимает кнопки на контроллере, выполняются соответствующие действия взаимодействия или захвата.

Скрипт намокания губки `SpongePaint` (рис. 5) также является классом в среде разработки *Unity*, который управляет взаимодействием губки с водой в контексте виртуальной реальности.

В классе определено поле `defaultMaterial`, которое хранит материал по умолчанию для объекта. При инициализации объекта метод `Start` получает текущий материал объекта и сохраняет его в `defaultMaterial`. Метод `resetMaterial` возвращает материал объекта к его состоянию по умолчанию, используя сохраненный ранее материал `defaultMaterial`.

Метод `OnTriggerEnter` вызывается, когда другой коллайдер входит в триггер. Если тег другого коллайдера равен «Water», то вызывается метод `Change()` у компонента `ChangeMaterialScript` этого объекта, если он существует. Затем тег для взаимодействия в `ToolActionToolScript` изменяется на «InteractsWithWetSponge».

То есть, этот скрипт позволяет губке взаимодействовать с водой, меняя свой материал и обновляя свои свойства для взаимодействия.

Заключение. Проект по созданию *VR*-приложения для реставрации культурных артефактов успешно реализован. Использование *Unity* и *C#* позволило создать интерактивное приложение, в котором пользователи могут восстанавливать и исследовать различные культурные артефакты. Программа *Blender* использовалась для создания точных *3D*-моделей артефактов, обеспечивая реалистичное и образовательное взаимодействие.

```
main.cs
1  using System.Collections;
2  using System.Collections.Generic;
3  using UnityEngine;
4
5  public class SpongePaint : MonoBehaviour
6  {
7      Material defaultMaterial;
8      private void Start()
9      {
10         defaultMaterial = GetComponentInChildren<MeshRenderer>
11         ().material;
12     }
13     public void resetMaterial()
14     {
15         GetComponentInChildren<MeshRenderer>().material =
16         defaultMaterial;
17     }
18     // Start is called before the first frame update
19     private void OnTriggerEnter(Collider other)
20     {
21         if (other.tag == "Water")
22         {
23             GetComponentInChildren<ChangeMaterialScript>
24             ().Change();
25             GetComponent<ToolActionToolScript>
26             ().tagToInteract = "InteractsWithWetSponge";
27         }
28     }
29 }
```

Рисунок 5. Скрипт намокания губки

Пользователи могут взаимодействовать с артефактами способами, недоступными в традиционной музейной среде, что способствует популяризации сохранения культурного наследия. Несмотря на успешность проекта, существуют области для дальнейшего улучшения, включая применение более сложных методов *3D*-моделирования и анимации. В целом, проект демонстрирует значительный прогресс в использовании *VR* для сохранения культурного наследия.

Список литературы

- [1] Sun, Y., Zheng, Y. (2022). A Study of Visitor Interaction with Virtual Museum. In: Stephanidis, C.,

Antona, M., Ntoa, S., Salvendy, G. (eds) HCI International 2022 – Late Breaking Posters. HCII 2022. Communications in Computer and Information Science, vol 1654. Springer, Cham. https://doi.org/10.1007/978-3-031-19679-9_44

[2] Wu, X., Chen, X., Zhao, J. et al. Influences of design and knowledge type of interactive virtual museums on learning outcomes: An eye-tracking evidence-based study. Educ Inf Technol (2023). <https://doi.org/10.1007/s10639-023-12061-4>

[3] Southall E, Hulusic V, Hargood C (2023) Interaction Styles in a Multi-Modal Virtual Museum. <https://doi.org/10.2312/GCH.20231183>

[4] Bekele MK, Champion E (2019) A Comparison of Immersive Realities and Interaction Methods: Cultural Learning in Virtual Heritage. Front Robot AI. <https://doi.org/10.3389/frobt.2019.00091>

[5] How galleries are going virtual | Apollo Magazine (apollo-magazine.com)

[6] Italy: share of museums using VR technologies 2020 | Statista

[7] Virtual Reality [VR] Market Size, Growth, Share | Report, 2030 (fortunebusinessinsights.com)

[8] DHQ: Digital Humanities Quarterly: Virtual museums as an extended museum experience: Challenges and impacts for museology, digital humanities, museums and visitors – in times of (Coronavirus) crisis

[9] How Art Galleries, Museums, & Exhibitions can benefit from a Virtual Tour (theredmarker.com)

[10] Unity User Manual 2021.3 (LTS) Unity - Manual: Unity User Manual 2021.3 (LTS) (unity3d.com).

[11] C# docs - get started, tutorials, reference. | Microsoft Learn. Режим доступа: <https://learn.microsoft.com/en-us/dotnet/csharp/>. – Дата доступа: 02.02.2024.

Авторский вклад

Авторы внесли равноценный вклад.

VIRTUAL REALITY APPLICATION «MUSEUM TOUR»

M.N. Melnikov

Student of the National Children's Technopark, student of the Gimnanisum No. 1 of Soligorsk

I.D. Polegin

Student of the National Children's Technopark, student of the Gimnanisum No. 10 of Gomel

P.V. Usenko

Software engineer, Department of Engineering Psychology and Ergonomics BSUIR, M.Sc.

M.S. Ilyasova

Assistant of the Department of Engineering Psychology and Ergonomics of BSUIR, M.Sc.

L.R. Korokin

Head of the Virtual and Augmented Reality Laboratory at the National Children's Technopark, Master of Technical Sciences

A.M. Prudnik

Associate Professor of Engineering Psychology and Ergonomics Department of BSUIR, Ph.D., Associate Professor

Abstract. The paper explores the impact of information technology on the interaction with cultural artifacts, focusing on virtual museums as a new stage in digital transformation. Various types of interactions between visitors and virtual objects are described, as well as their contribution to enhancing visitor experiences and understanding of cultural heritage. The paper discusses development tools for VR applications, including Unity, C#, and Blender, and their role in creating an interactive application. The process of implementing a VR application is described, which includes the selection of an artifact for restoration, familiarization with tools and materials, the application of various restoration methods, and the creation of an endless cycle of learning. The paper emphasizes that the application promotes user engagement and motivation, learning, self-expression, and creativity, as well as improving skills and knowledge in the field of restoration. The use of Unity engine's graphic resources ensures a high level of visual realism and the effect of full immersion. The successful implementation of the project confirms the effectiveness of the VR application for the restoration of cultural artifacts as a tool for learning and developing skills in the field of restoration, and also promotes the popularization of cultural heritage preservation.

Keywords: Virtual reality, Unity, C#, object-oriented programming, 3D modeling, cross-platform development, interactive environment, real-time rendering, restoration of cultural artifacts, education in the field of cultural heritage, visualization of historical context.

УДК 004.93

СЛЕДУЕТ ЛИ НОРМАЛИЗОВАТЬ ДАННЫЕ ДИНАМИЧЕСКИХ ПОДПИСЕЙ ПЕРЕД ВЕРИФИКАЦИЕЙ МЕТОДОМ DTW?



В.В. Старовойтов

Главный научный сотрудник ОИПИ
НАН Беларуси,
доктор технических наук, профессор
valerys@newman.bas-net.by

В.В. Старовойтов

Окончил Белорусский государственный университет. Лауреат Государственной Премии Республики Беларусь (2002г.). Область научных интересов связана с анализом цифровых изображений.

Аннотация. В статье исследуются динамические подписи, выполненные человеком на планшете. Рассматривается задача нормализации данных на этапе предварительной обработки перед проверкой подлинности цифровой подписи на базе метода динамической трансформации временной шкалы (*dtw*). Приведены результаты экспериментов, выполненных на данных из самой большой общедоступной базы динамических подписей *DeepSignDB*. Впервые экспериментально доказано, что вопреки общепринятой практике расстояния между ненормализованными параметрами подписи и их производными, вычисленные методом *dtw*, позволяют точнее разделять поддельные подлинные подписи чем нормализованные по амплитуде. Это позволяет точнее и объективнее учитывать индивидуальные особенности выполнения динамической подписи конкретным человеком, что особенно важно при выполнении экспертизы при наличии ограниченного числа доступных для сравнения подлинных подписей.

Ключевые слова: верификация, динамическая подпись, *dtw*-преобразование, параметрические функции.

Введение. Подпись выполняет две функции: удостоверяет личность подписавшего и заверяет подписанный им документ. С точки зрения информационных технологий распознавание подписи относится к биометрическим технологиям, которые различны при анализе подписей, выполненных разными способами: статически (на бумаге) и динамически или онлайн (на планшете). Динамическая подпись содержит данные о динамике исполнения подписи, которые невидимы и делают подпись более уникальной, они точнее характеризуют особенности исполнения подписи конкретным человеком и лучше защищены от подделок. Не стоит путать динамическую подпись с электронной цифровой подписью (ЭЦП), введенной в нашей стране законом республики Беларусь от 28 декабря 2009 г. N 113-3 «Об электронном документе и электронной цифровой подписи». Основным элементом ЭЦП является криптографический ключ. Этот тип подписи используют в основном юридические лица. С 3 марта 1918г. постановлением Правления Национального банка Республики Беларусь [1] разрешена к использованию цифровая рукописная подпись. Это и есть динамическая подпись. Согласно постановлению, цифровая рукописная подпись – это собственноручная подпись клиента, учиненная с помощью соответствующих программных средств (в том числе планшетов) для подтверждения целостности и подлинности подписываемого документа в электронном виде. Официальных методик исследования экспертами подлинности таких

подписей в Республике Беларусь до настоящего времени нет, а в России отсутствует даже официально принятое понятие цифровой рукописной подписи.

Как указано в работе [2] «система признаков, используемых на данный момент экспертами-почерковедами, была разработана еще в советское время, и на сегодняшний день не особо претерпела существенных изменений». В работе [3] отмечено, что «методические основы отечественной почерковедческой экспертизы заложены в 60-70-х годах прошлого столетия». «В настоящее время общая методика проведения почерковедческого исследования особых изменений не претерпела». «Используя одну и ту же методику проведения идентификационной почерковедческой экспертизы, разные эксперты могут по-разному оценить выявленные совпадения и различия. В результате по одному и тому же исследуемому объекту могут быть сделаны совершенно противоположные (иногда категоричные) выводы. Во многом это связано с тем, что используемые идентификационные признаки в большинстве своем носят качественный, оценочный характер и формируются на основе субъективной оценки эксперта».

Эксперты-юристы все еще используют визуальный анализ графического представления цифровых подписей. Приведенные факты свидетельствуют об актуальности разработки интеллектуальных систем проверки подлинности динамических подписей для повышения объективности анализа, выполняемого экспертом. Во многих странах ведутся активные исследования направленные на поиск решения задачи верификации (проверки подлинности) динамической подписи [4].

Представление динамических подписей. Подпись, зарегистрированная на планшете с помощью специального стилуса, представляют собой несколько дискретных параметрических функций. Обязательными из них являются: координаты X и Y положения стилуса, время фиксации этих координат T и давление стилуса на поверхность планшета P в каждой точке. Точки нахождения кончика стилуса фиксируются через равные промежутки времени. У разных производителей они составляют 5, 10 или 15 мсек. Некоторые типы планшетов дополнительно регистрируют две угловые координаты положения стилуса. Все параметры, обычно, представлены целыми числами в определенной шкале, в текстовом формате. Некоторые производители не сохраняют координаты стилуса, перемещаемого без давления на планшет, а только отмечают нулями места разрывов дискретных функций. На рисунке 1 приведены примеры визуального представления подлинной и поддельной подписей данного типа. При визуализации подписи координаты точек касания стилусом планшета соединяются отрезками прямых и формируется ее привычное представление. Некоторые исследователи дополнительно анализируют «невидимые» траектории стилуса при выполнении подписи (на рисунке 1 они показаны красным цветом).

При анализе подлинности цифровых подписей в экспертной практике чаще используется их представление в виде дискретных параметрических функций X , Y , P (рисунок 2). Функции являются параметрическими, поскольку их значения каждой из них зависят только от момента регистрации t .

Постановка задачи. Задачу верификации динамической подписи, ориентированную на практическое применение можно сформулировать следующим образом. Дано N подлинных динамических подписей некоторого человека и одна подпись, подлежащая проверке. Требуется разработать метод формирования объективной оценки сходства верифицируемой подписи с подлинными. Для этого необходимо оценить влияние нормализации сравниваемых данных на результаты верификации подписей человека.

Авторы статьи [5] выполнили исследования по анализу зависимости результатов dtw -преобразований от пространственной нормализации данных (приведения функций к фиксированному числу значений), примененных к данным из 4-х разных баз данных. Они доказали, что двусторонний тест с уровнем значимости 0,05 между парами

параметрических функций переменной длины и равной длины показывает, что нет статистически значимой разницы между точностью вычисления *dtw*-расстояний.

В работе [6] сравнивались четыре различных типа нормализации данных динамических подписей: 1) отсутствие нормализации; 2) нормализация по времени; 3) масштабирование амплитуды давления P до диапазона $[0; 1]$; и 4) нормализация по времени с последующей нормализацией давления P . Нормализация была применена до выполнения *DTW*. Авторы выделили 15 лучших вариантов нормализованных признаков с наименьшими ошибками верификации *EER*. Почти все из 15 вариантов связаны со скоростью.

Основные подходы к решению задачи верификации динамических подписей.

Основные методы верификации подписей, разрабатываемые специалистами в области ИИ, описаны в ряде обзоров. Последний из них представлен в статье [4]. Подходы на базе ИИ используют метод ближайшего соседа, скрытые марковские модели, машины опорных векторов, нейросетевые методы, нечеткие классификаторы и другие. Популярно применение преобразования называемого динамической трансформацией временной шкалы или *DTW*-преобразования. Его разработал Беллман в 1952г. [7]. С 1977г. это преобразование все активнее стало использоваться для сравнения динамических подписей [8].

Недостаток применения *HMM* для анализа динамической подписи заключается в том, что для этого требуется вычисление большого числа признаков, объем данных для обучения модели должен быть очень большим и высока временная сложность [9]. Для обучения искусственных нейронных сетей также требуется большое число образцов подлинных и поддельных подписей. Следует отметить, что на практике у эксперта поддельных динамических подписей не бывает. В лучшем случае он имеет несколько подлинных подписей, определенно выполненных неким человеком и верифицируемую подпись, подлинность которой требуется установить. Указанные ограничения затрудняют разработку классификаторов, указанных типов.

Активно исследуются методы распознавания фальшивых подписей, оценивающих сходство кривых, описываемых параметрическими функциями. Многие из них базируются на оценке расстояния между двумя кривыми. Здесь возможны 2 варианта: вычисление максимального расстояния между совмещенными кривыми (расстояние Фреше [10]) и наименьшее суммарное расстояние необходимое для совмещения двух кривых на базе *dtw*-преобразования. Первый чувствителен к выбросам, т.к. результатом фактически является наибольшее расстояние между всеми возможными парами точек двух кривых.

Dtw-преобразование вычисляет интегральные расстояния между двумя дискретными параметрическими функциями X , Y , P , либо производными от них функциями. Поэтому оно объективнее оценивает сходство кривых. Многие исследователи используют нормализацию данных при верификации подписей посредством *dtw*. Однако в литературе описаны только сравнительные исследования по нормализации длины указанных функций [5-6], но не амплитуды. В данной работе экспериментально сравнивается возможность разделения множества *dtw*-расстояний, вычисленных между параметрическими кривыми подлинных и поддельных подписей одного человека, что в итоге позволяет определить поддельные подписи.

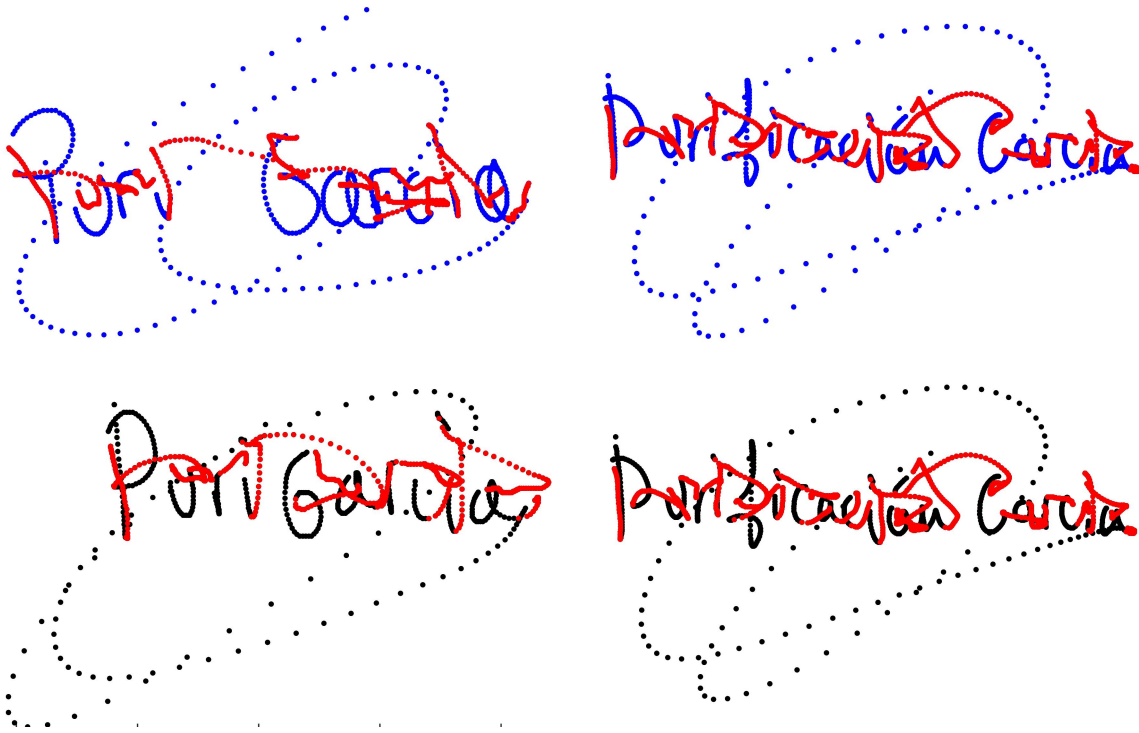


Рисунок 1. В верхней строке представлены фальшивые подписи (левая верхняя – самая сложная для верификации), в нижней – подлинные. Красным цветом показаны точки, в которых стилус был поднят

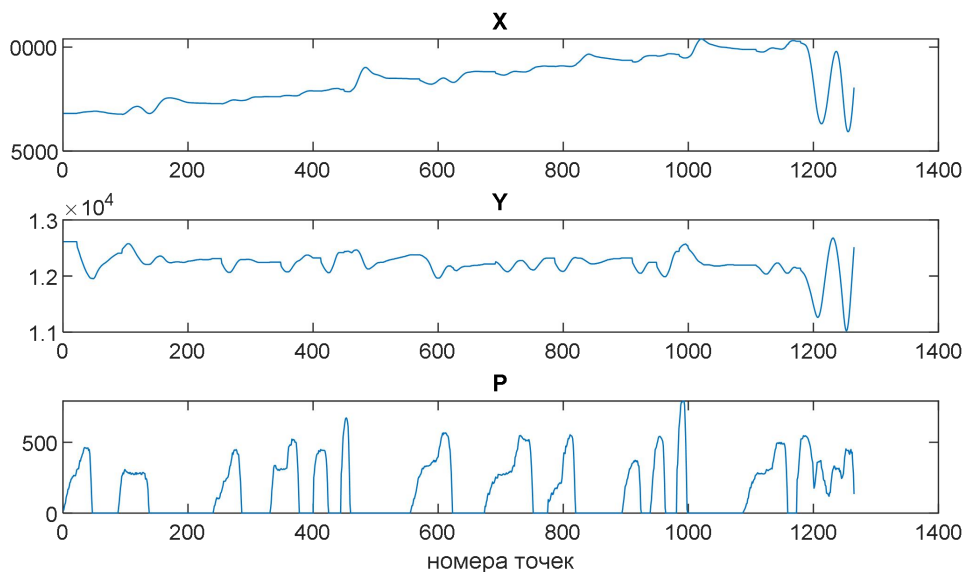


Рисунок 2. Основные данные о динамической подписи в виде дискретных параметрических функций

Полученные результаты могут быть использованы в разработке метода верификации при ограниченном числе подлинных динамических подписей и отсутствии у эксперта поддельных.

Преобразование DTW. Кратко опишем идею преобразования *dtw*. Имеются две параметрические функции, представленные наборами чисел разной длины и амплитуды (a_1, a_2, \dots, a_n) и (b_1, b_2, \dots, b_m). На первом этапе вычисляются локальные расстояния между элементами двух последовательностей, формируется матрица таких расстояний d_{ij} из n

строк и m столбцов. На следующем этапе вычисляются минимальные расстояния в матрице между первым элементом с индексами $i=1, j=1$ и последним элементом с индексами $i=n, j=m$ при помощи алгоритма динамического программирования и критерия оптимизации:

$$a_{ij} = d_{ij} + \min(a_{i-1,j-1}, a_{i-1,j}, a_{i,j-1}),$$

где d_{ij} – минимальное расстояние между первым элементом матрицы с индексами 11 и ij .

Результатом преобразования является кратчайший путь (и его длина) сдвига точек одной последовательности к точкам другой, который можно изобразить набором отрезков на рисунке 3.

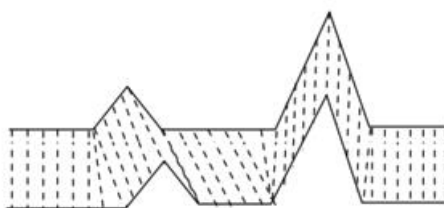


Рисунок 3. Иллюстрация идеи преобразования dtw

Применительно к динамическим подписям это можно проиллюстрировать на рисунке 4. Множества y -координат двух подписей нормализованы в диапазон $[0; 1]$ и показаны на верхнем рисунке. На нижнем рисунке показан результат совмещения этих кривых на базе преобразования dtw и показано расстояние между исходными кривыми. Оно равно 14.60. Чем меньше расстояние, тем больше одна кривая совпадает с другой.

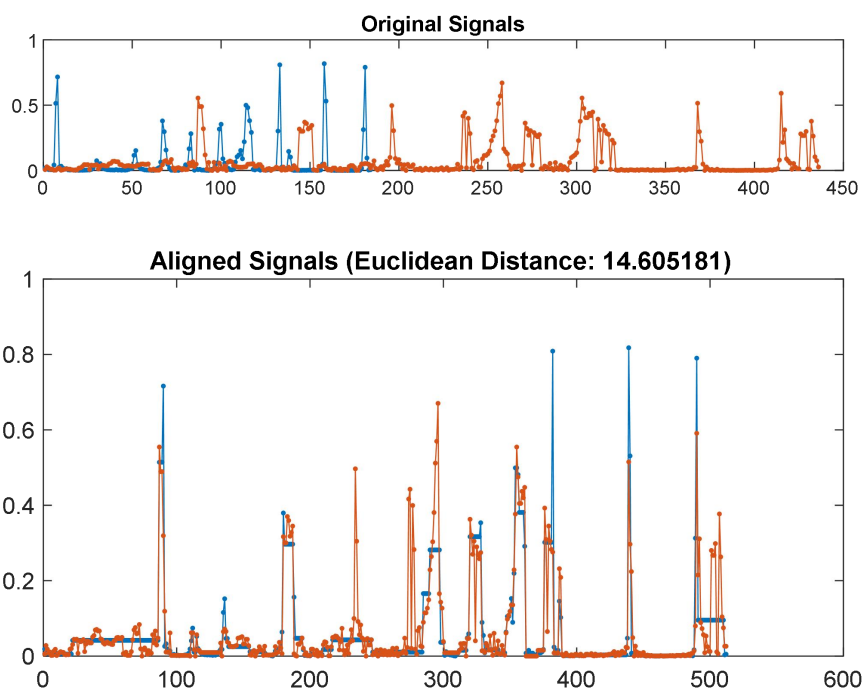


Рисунок 4. Совмещение двух параметрически заданных функций с помощью преобразования dtw

Если кривые имеют разные диапазоны значений, расстояние между ними будет больше, чем между нормализованными и результирующая (после совмещения) дискретная кривая будет иметь больше элементов. На рисунках 5 и 6 представлены пары синусоид с одинаковым сдвигом, но разными амплитудами. Расстояние, вычисленное в результате применения *dtw*, увеличивается в 47 раз и в результате совмещения число точек кривой увеличивается в 1.5 раза, т.е. кривые на рисунке 5 больше похожи друг на друга, чем на рисунке 6. Результат совмещения кривых оказывается длиннее исходных кривых. Этот пример показывает, что для сравнения сходства двух кривых равной длины и одинаковой формы их следует нормализовать к единому диапазону значений. В соответствии с этими рекомендациями в большинстве исследований по верификации динамических подписей выполняют их нормализацию.

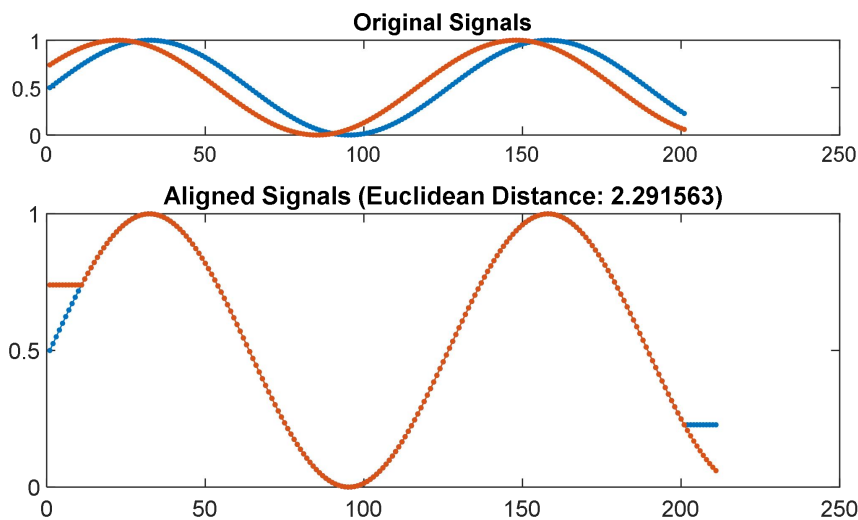


Рисунок 5. Совмещение двух синусоид с одинаковыми амплитудами, но со сдвигом в 0.5 радиана

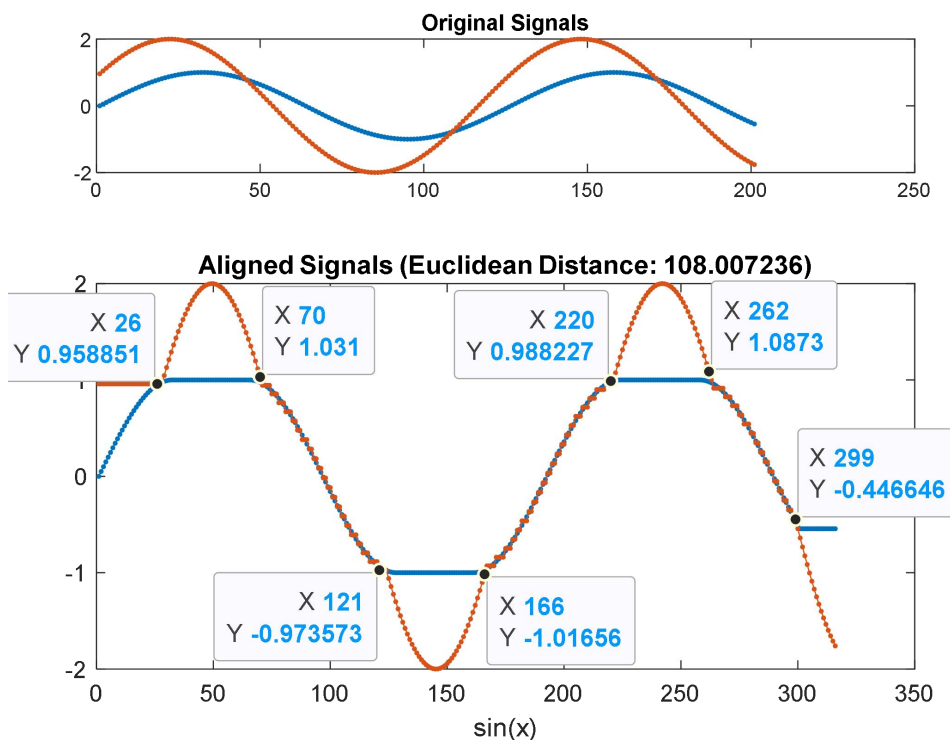


Рисунок 6. Совмещение тех же кривых, но у одной амплитуда в 2 раза выше

Экспериментальные исследования. Эксперименты выполнялись на данных самой большой доступной базе динамических подписей *DeepSignDB* [11]. Эта база содержит более 40 000 динамических подписей 1526 человек, собранных из 5 других баз подписей и зарегистрированных на планшетах 8 типов.

При выполнении экспериментов случайным образом выбрано 10 человек. Для каждого случайно выбиралось по 7 подлинных и 7 поддельных подписей. Все подлинные подписи одного человека сравнивались между собой (21 пара для каждого человека), затем оценивалась сходство пар (подлинная-фальшивая) подписей этого же человека из числа выбранных, всего 49 пар для каждого. Сравнение выполнялось посредством совмещения дискретных параметрических кривых, описывающих локальные изменения скорости и ускорения при начертании подписей методом *dtw*, вычисляющим расстояние между парами кривыми. Чем меньше расстояние, тем больше кривые похожи друг на друга. С другой стороны, эти расстояния могут служить признаками, позволяющими разделять множества пар подлинных подписей от множества с поддельными подписями. Возможность порогового разделения важнее абсолютных значений этих признаков.

Исследовалась гипотеза, согласно которой человек выполняет свои подлинные подписи быстрее чем поддельные, вместо другого человека. Предполагалось, что расстояние между точками, фиксируемыми через равные промежутки времени, в подлинных подписях будет больше. Это оказалось верным не во всех случаях. Кроме этого, подпись в основном представляет собой краткую запись, которая обычно выполняется слева направо (за исключением арабских стран и Израиля, но это не принципиально). Основные вариации подписи при написании происходят в вертикальном направлении, а менее значительные в горизонтальном. Зависимости в изменении давления P на стилус в подлинных подписях одного человека не замечено. Поэтому исследовались данные, о скорости и ускорении, вычисленные по параметрическим функциям X и Y каждой подписи.

Сравнение выполнялось посредством оценки изменений скорости и ускорения нанесения подписи в горизонтальном, вертикальном направлении и абсолютных направлениях в плоскости XU . Первые производные далее обозначены dx , dy , dxu , вторые – ddx , ddy , $ddxu$). Они вычислялись через разности значений смежных точек. Для 7 подлинных и 7 поддельных подписей каждого из 10 человека были вычислены расстояния между 21 парой подлинных подписей и 49 парами поддельных с подлинными. Общее число пар составило 700. Для каждой пары вычислялось 10 оценок, указанных в таблице 1.

В таблице 1 представлено количество ошибок при сравнении *dtw*-расстояний между 49 парами подлинных и поддельных подписей относительно максимальных расстояний между 21 парой из подлинных подписей для каждого из 10 человек, т.е. число ошибок при пороговом отделении поддельных подписей посредством изменений скоростей и ускорений в начертании подписей. Жирным шрифтом выделено наименьшее число ошибок по каждому человеку. Используя вычисленные расстояния как признаки пары с поддельной подписью в большинстве случаев можно точнее установить порог индивидуально для каждого человека и каждого параметра. Объединяя разные параметры в признаковое пространство и настраивая на этапе обучения персонально для каждого человека пороги, можно строить индивидуальные деревья решений для верификации динамических подписей конкретного человека.

На представленных ниже рисунках 7-10 показаны соотношения *dtw*-расстояний между парами кривых, представляющих отдельные характеристики сравниваемых подписей. На каждом рисунке представлено 21 значение для подлинных подписей (показаны синими звездочками) и 49 значений расстояний между поддельными и

подлинными подписями (показаны красными кружками). По горизонтали указаны номера пар, по вертикали – расстояния.

На рисунке 7 и 8 показаны оценки для ненормализованного и нормализованного представления данных. Оба варианта позволяют отделить поддельные подписи 330-го человека по одному признаку – вертикальная скорость. Однако, можно заметить изменения соотношений между оценками соседних пар подписей. Это свидетельствует о том, что нормализация данных по амплитуде может повлиять на результат верификации.

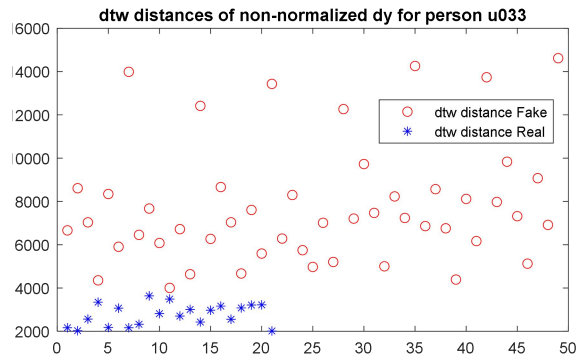


Рисунок 7. Оценки после *dtw*-преобразования 21 пары подлинных подписей и 49 пар поддельных с подлинными. Сравнялись скорости изменения написания подписи человеком № 330 в вертикальной плоскости, Данные не нормализованы

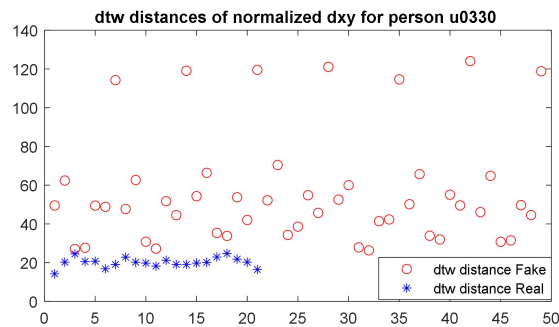


Рисунок 8. Такие же оценки для тех же данных после нормализации

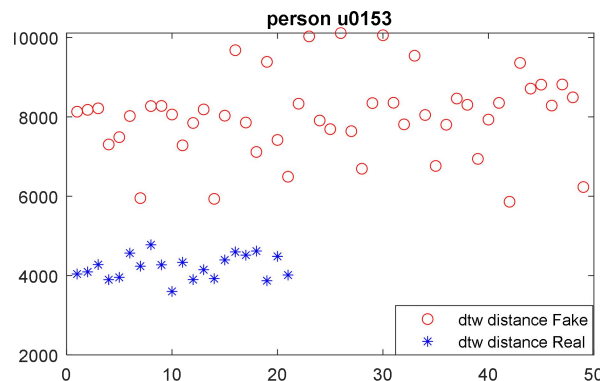


Рисунок 9. Расстояния после *dtw*-преобразования пар ненормализованных подписей человека № 153 позволяют отделить пары с поддельными подписями

Рисунок 9 демонстрирует, что ненормализованные данные позволяют отделить поддельные подписи человека № 153 пороговым методом, а рисунок 10 – что эти же нормализованные данные не позволяют. Этот факт говорит о том, что особенности

нанесения подписи на планшет человеком №153 существенно отличаются от нанесения поддельных подписей другими людьми, а нормализация данных сглаживает различия.

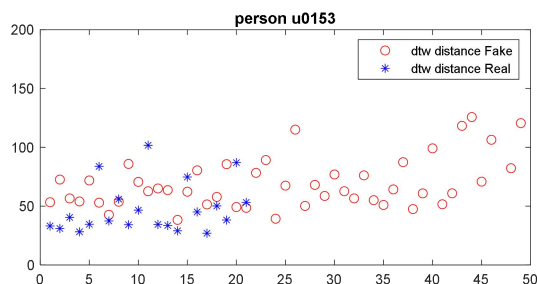


Рисунок 10. После нормализации тех же данных и выполнения *dtw*-преобразований значения расстояний между двумя группами подписей перемешиваются

Таблица 1. Число ошибок при верификации 49 пар подлинных и поддельных подписей относительно максимальных расстояний между парами подлинных

№	человек	Для ненормализованных данных X, Y					Для нормализованных данных X, Y				
		dx	dy	dxy	ddy	$ddxy$	dx	dy	dxy	ddy	$ddxy$
1	u0330	1	0	0	0	2	20	3	0	32	15
2	u0336	10	28	11	35	25	44	47	11	45	43
3	U0322	16	0	0	0	1	11	0	2	46	24
4	U0329	0	0	0	12	18	41	46	11	43	34
5	U0222	9	9	13	19	16	40	31	46	43	41
6	U0153	1	0	0	3	15	26	45	31	44	15
7	U0497	7	21	4	4	0	13	15	12	29	10
8	U0250	0	7	3	7	5	1	7	5	9	3
9	U0050	1	3	0	17	6	12	6	2	48	21
10	U0001	2	3	4	4	11	9	43	15	48	16

В результате анализа экспериментальных данных выявлены следующие факты.

1 Согласно расстояниям, вычисленным *dtw*-преобразованиями иногда нормализованные поддельные подписи больше похожи на подлинные подписи человека, чем не нормализованные.

2 Подлинные подписи человека под номером 336, представленные на рисунке 1, сложны для анализа, поскольку они имеют существенно различающий их фрагмент – ряд букв в центре подписи (на некоторых подписях эти буквы отсутствуют). Это приводит к тому, что при сравнении подлинных, но существенно различных подписей оценки их сходства хуже, чем при сравнении некоторых поддельных с отдельными подлинными. При выборе подлинных подписей для выполнения верификации следует избегать таких образцов.

3 При оценке принадлежности подписи конкретному человеку важны не абсолютные значения расстояний, вычисленных методом *dtw*, а возможность разделения подлинных и поддельных подписей.

Заключение. В работе на примерах анализа реальных динамических подписей разных людей показано, что при верификации подписей такого типа при использовании метода динамической трансформаций временной шкалы (*DTW*-преобразования) важнее не минимальность абсолютных значений расстояний, вычисленных этим методом, а возможность разделения подлинных и поддельных подписей одного человека, опираясь

на вычисленные значения. Min-max нормализация амплитуды параметрических функций, описывающих данные динамических подписей, может снизить точность объективной верификации таких подписей. При использовании *dtw*-преобразования нормализацию амплитуды данных о динамической подписи выполнять не рекомендуется.

Список литературы

- [1] Цифровая рукописная подпись как новый способ подтверждения документа в электронном виде вводится в Беларуси. [Электронный ресурс]. – Режим доступа: <https://pravo.by/novosti/novosti-pravo-by/2018/march/27952/>.
- [2] Ярощук И.А., Гриневич К.В., Проблемные вопросы экспертизы подписи как малообъемного почеркового объекта // Актуальные проблемы российского права.- 2021.- Т. 129.- №.8.- С.141-151.
- [3] Мещеряков В.А., Бутов В.В., Оценка возможностей почерковедческой экспертизы сквозь призму современных информационных технологий. Вестник Воронежского института МВД России. 2017ю- № 2.- С.40-46.
- [4] Kaur H., Kumar M., Signature identification and verification techniques: state-of-the-art work, Journal of Ambient Intelligence and Humanized Computing, 2023, Vol. 14, No.2, pp.1027-45.
- [5] Ratanamahatana CA, Keogh E. Everything you know about dynamic time warping is wrong. In 3rd workshop on mining temporal and sequential data 2004 Aug 22, Seattle, USA, pp.50-60.
- [6] Fenton, D., Bouchard, M., & Yeap, T. H. (n.d.). Evaluation of Features and Normalization Techniques for Signature Verification Using Dynamic Timewarping. 2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings. doi:10.1109/icassp.2006.1660860.
- [7] Bellman R. On the theory of dynamic programming. Proceedings of the national Academy of Sciences. 1952, Vol. 38(8), pp.716-719.
- [8] Yasuara M., Oka M., Signature verification experiment based on non-linear time alignment, IEEE Trans. on Systems, Man and Cybernetics, 1977, Vol.17, pp. 212-216.
- [9] Chughtai J. R., Khalid S. Online Signature Verification: A Review, Journal Appl. Environ. Biol. Sci., 2014, (9S), pp.303-308.
- [10] Zheng J, Gao X, Zhan E, Huang Z. Algorithm of on-line handwriting signature verification based on discrete Fréchet distance. In 3rd Int. Symposium on Advances in Computation and Intelligence, Wuhan, China, December 19-21, 2008, pp. 461-469.
- [11] DeepSignDB [Электронный ресурс]. – Режим доступа: <https://github.com/BiDALab/DeepSignDB>.

Авторский вклад

Старовойтов Валерий Васильевич – выполнил все эксперименты и написал статью.

SHOULD WE NORMALIZE DYNAMIC SIGNATURES DATA BEFORE DTW-BASED VERIFICATION?

V.V. Starovoitov

Chief Researcher UIIP NAS

of the Republic of Belarus

Doctor of Engineering Sciences, Professor

Abstract. The paper examines dynamic signatures made by a person on a tablet. We consider the problem of data normalization at the preprocessing stage before verifying the authenticity of a digital signature based on the dynamic timeline warping (*dtw*) method. The results of experiments performed on data from the largest public database of dynamic signatures, DeepSignDB, are presented. For the first time, it has been experimentally proven that, contrary to generally accepted practice, the distances between non-normalized signature parameters and their derivatives, calculated by the *dtw* method, make it possible to more accurately separate fake signatures than those normalized by amplitude. This makes it possible to more accurately and objectively take into account the individual characteristics of the execution of a dynamic signature by a specific person, which is especially important when performing an examination when there is a limited number of authentic signatures available for comparison.

Key words: verification, on-line signature, *dtw* transformation, parametric functions.

УДК 004.891

ОБУЧЕНИЕ ВОПРОСНО-ОТВЕТНОЙ НЕЙРОСЕТЕВОЙ МОДЕЛИ НА БАЗЕ АРХИТЕКТУРЫ МОДЕЛИ LLaVA 1.5 С ЭНКОДЕРОМ SAIGA MISTRAL 7B И АЛГОРИТМА НИЗКОРАНГОВОЙ АДАПТАЦИИ LORA



Л.А. Демидова

д.т.н., профессор кафедры корпоративных информационных систем Института информационных технологий МИРЭА – Российского технологического университета, Москва, Россия;
orcid.org/0000-0003-4516-3746,
e-mail: demidova.liliya@gmail.com



Н.А. Морошкин

аспирант кафедры корпоративных информационных систем Института информационных технологий МИРЭА – Российского технологического университета, Москва, Россия;
orcid.org/0009-0002-8787-2452, e-mail: moroshkin@mirea.ru

Л.А. Демидова

Окончила Московский государственный университет им. М.В. Ломоносова (МГУ). Область научных интересов связана с разработкой методов и алгоритмов интеллектуального анализа данных.

Н.А. Морошкин

Окончил МИРЭА – Российский технологический университет. Область научных интересов связана с разработкой вопросно-ответных нейросетевых моделей компьютерного зрения и методов оптимизации, конвертации нейросетевых моделей.

Аннотация. В данной работе реализован алгоритм обучения визуальной вопросно-ответной нейросетевой модели на базе архитектуры нейросетевой модели LLaVA 1.5 с использованием текстового энкодера Mistral 7b, позволяющего улучшить результаты работы модели в задаче визуального вопросно-ответного моделирования и алгоритма низкоранговой адаптации LoRA, позволяющего ускорить процесс обучения модели.

Показано, что задача вопросно-ответного моделирования может быть решена нейросетевыми моделями с использованием больших языковых моделей, описана методика ускорения обучения таких моделей. Проведена оценка эффективности разработанной модели и показаны общие аспекты обучения вопросно-ответных моделей на наиболее популярных вопросно-ответных наборах данных.

Ключевые слова: Визуальное вопросно-ответное моделирование, большие языковые модели, алгоритм низкоранговой адаптации LoRA, нейросетевая модель LLaVA 1.5, Saiga/Mistral 7b.

Введение. Вопросно-ответное моделирование (BOM, Question Answering, QA) давно и широко применяется при решении различных прикладных задач [1]. Однако, начиная с 2017 года [2], специалисты-разработчики начали проявлять интерес к синтезу вопросно-ответных систем (ВОС) с использованием технологий искусственного интеллекта, а именно – с применением нейросетевого моделирования. При этом особое внимание стало уделяться визуальному вопросно-ответному моделированию (BBOM, Visual Question

Answering, VQA), которое обычно предполагает решение мультимодальной задачи обработки больших данных.

Преимущество такого подхода заключается в возможности адаптации ВОС к решению множества задач визуального анализа данных, например, к детектированию объектов на визуальных данных, классификации визуальных объектов. Такие подходы реализуются с помощью создания текстовых запросов к нейросетевым моделям [1].

В последних работах [3,4], посвященных обработке текстовых данных, наиболее популярное решение является использование больших языковых моделей. Подобные модели позволяют решать множество задач обработки естественного языка [4]. Задача визуального вопросно-ответного моделирования – это мультимодальная задача, так как она содержит в себе объединение подходов к обработке визуальных и текстовых данных. Вопрос применения больших языковых моделей в качестве текстовых энкодеров в решениях визуального вопросно-ответного моделирования исследован в меньшей степени, однако исследователи нейросетевых технологий начинают использовать инновации больших языковых моделей для построения больших мультимодальных моделей [5].

Визуальная вопросно-ответная нейросетевая модель LLaVA 1.5. Нейросетевая модель LLaVA 1.5 – это большая мультимодальная модель, решающая задачу вопросно-ответного моделирования с использованием большой языковой модели в качестве энкодера текстовых данных [5]. Преимущества данной архитектуры заключаются в открытом исходном коде, в возможности использования множества известных больших языковых моделей, а также – в возможности решения задачи распознавания символов на изображениях.

Архитектура модели представлена на рисунке 1. Ядро модели составляют большая языковая модель *Vicuna-7b* (с 7 миллиардами параметров) и визуальный энкодер *CLIP ViT-L/14* с 428 миллионами параметров.

Ядро алгоритма работы модели LLaVA состоит из соединения двух энкодеров – визуального и текстового. Для входной многомерной матрицы изображения X_v используется визуальный энкодер (в оригинальной архитектуре версии 1.5 нейросетевая модель *CLIP ViT-L/14*), который обрабатывает изображение, получая векторное представление Z_v . Затем векторное представление Z_v передается в многослойный перцептрон, размерность выходного слоя которого совпадает с размерностью входного слоя большой языковой модели (в оригинальной архитектуре версии 1.5 используется большая языковая модель *Vicuna*), получая векторное представление изображения H_v . Текстовый запрос к модели X_q (вопрос к вопросно-ответной системе) токенизируется с помощью алгоритма токенизации. Под токенизацией в контексте этой работы имеется в виду алгоритм разделения текста на элементарные составляющие (символы или слова). В результате формируется токенизированное представление H_q , размерность которого равна размерности входного слоя большой языковой модели. Наконец, векторные представления изображения H_v и текстового запроса H_q передаются в языковую модель. В результате формируется текстовый ответ X_a .

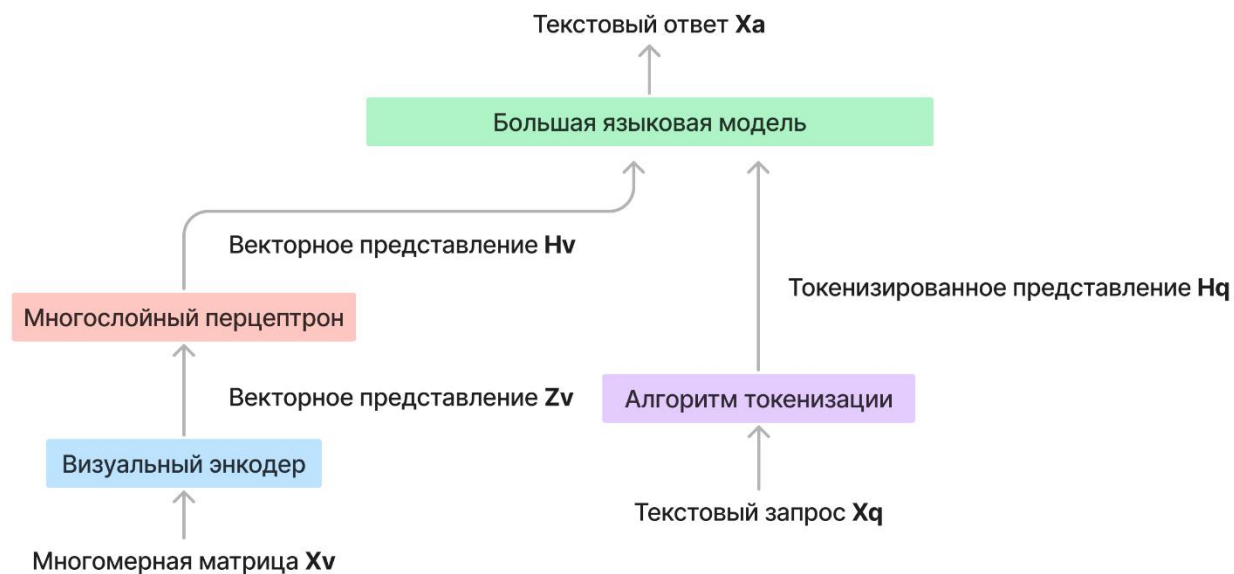


Рисунок 1. Архитектура вопросно-ответной нейросетевой модели LLaVA 1.5

Особенность модели LLaVA заключается в возможности обучения модели без использования больших наборов данных по сравнению с другими архитектурами [5].

Изменив стандартный подход к формированию запроса [5], модель обладает лучше обобщающей способностью и способна отвечать развернуто на запросы. Стандартный подход состоял в том, что модели задавался один или несколько вопросов к одному контексту (в задаче BBOM контекстом является входное изображение). В модели LLaVA было предложено использовать множественные вопросы к множественным контекстам. Тем самым, авторы улучшали способность к обобщению у модели.

Кроме того, изменен подход к обучению вопросно-ответной модели: в вопрос к ВОС добавлены инструкции, которые указывают большой мультимодальной модели на необходимость распознавания символов на изображениях для более точного ответа. Оценки точности модели LLaVA и некоторых других SOTA моделей представлены в таблице 1. В таблице представлены наиболее распространенные наборы данных для обучения и тестирования моделей BBOM.

Таблица 1. Оценки точности SOTA моделей при решении задачи VQA на разных наборах данных [6]

Имя модели	Набор данных VQA _{v2} , метрика accuracy [5]	Набор данных GQA, метрика accuracy [5]	Набор данных POPE, метрика accuracy [5]
LLaVA 1.5	0,8	0,63	0,85
InstructBLIP	0,48	0,49	0,79
BLIP-2	0,15	0,02	0,86

Авторы модели LLaVA уделили большое внимание возможности масштабирования входного изображения, улучшив обобщающую способность визуального энкодера. В следующей версии 1.6 идея масштабирования входного изображения была расширена, что позволило улучшить метрические показатели модели (f -мера [6], точность).

В качестве наборов обучающих данных для модели LLaVA 1.5 можно использовать стандартные наборы данных для BBOM, наборы данных визуального регионального представления (REVIVE, Revive Dataset) и наборы данных для решения задачи распознавания символов на изображениях (OCR, Optical Character Recognition).

Модель *LLaVA* характеризуется также способностью работать с несколькими языками. Несмотря на то, что модель изначально не была адаптирована под многоязыковую задачу, она способна обрабатывать запросы на не родных языках. Однако качество таких ответов хуже, чем у аналогичных моделей. Главные особенности модели, такие как региональное представление и возможность распознавать символы на изображениях, теряются из-за невозможности большой языковой модели обработать векторное представление неродных языков.

Текстовый энкодер *Saiga/Mistral 7b*. В данной работе предлагается обучить модель *LLaVA* на оригинальном наборе данных, заменив текстовый энкодер на модель *Saiga/Mistral 7b*, обученную на наборе данных на русском языке. Модель *Saiga/Mistral 7b* использует внимание к групповым запросам для ускорения работы и алгоритм скользящего окна внимания для обработки последовательности произвольной длины [7]. Учитывая, что оригинальная модель *Mistral* была обучена с применением групповых запросов и то, что модель *LLaVA* использует аналогичные данные для валидации и аннотации, можно сделать предположение о том, что модель *Saiga/Mistral 7b* способна заменить модель *Vicuna* в роли текстового энкодера в архитектуре *LLaVA*. Оценки точности модели *Mistral 7b* и её некоторых аналогов представлены в таблице 2. Учитывая специфику вопросно-ответного моделирования, сравнение больших текстовых моделей представлено в виде MT-Bench оценки, основанной на рейтинговой системе. Эта оценка показывает, насколько модель способна вести многоходовый диалог, начиная с одного из 80 подготовленных вопросов [7].

Таблица 2. Оценки точности больших языковых моделей для вопросно-ответного моделирования [7]

Имя модели	Набор данных <i>MT-Bench</i> [7]
<i>Mistral 7b Instruct</i>	6,84
<i>Llama 2 7b</i>	6,65
<i>Vicuna 13b</i>	6,17

Набор данных *Saiga*, использующийся для обучения модели *Saiga/Mistral 7b*, основан на принципе работы большой языковой модели *Baize* [8] как конвейера для генерации многоходового корпуса текстовых данных. Данный подход адаптирован под решение задачи вопросно-ответного моделирования [9] и демонстрирует хорошую обобщающую способность для ведения многоходового диалога.

Обучение вопросно-ответной модели с использованием алгоритма низкоранговой адаптации *LoRA*. При обучении вышеописанной визуальной вопросно-ответной модели предлагается использовать алгоритм низкоранговой адаптации *LoRA*, который в свою очередь применяется в процессе обучения оригинальной модели *LLaVA 1.5*, а также в некоторых других моделях для улучшения производительности обработки разных доменов корпуса текстов [6].

В основу алгоритма *LoRA* [5] закладывается разделение матрицы весов модели на две матрицы меньшего размера. Эти матрицы определяют новые весовые коэффициенты нейросетевой модели. Затем модель с новыми весовыми коэффициентами обучается стандартным методом [5]. Такой подход позволяет использовать меньше вычислительных ресурсов, так как две меньшие матрицы требуют ресурсов на хранение и обработку меньше, чем одна большая, составленная из двух малых. Алгоритм *LoRA* получил широкое распространение в исследованиях больших языковых моделей [7].

При проектировании архитектуры решения было необходимо внести изменения в оригинальную архитектуру, поскольку размер векторного представления оригинальной модели *Vicuna 13b* и модели *Saiga/Mistral 7b* отличаются: они равны 5120 выходных

нейронов и 4096 выходных нейронов соответственно. Для дальнейшего обучения необходимо модифицировать архитектуру многослойного персептрона, изменив количество выходных нейронов с 5120 на 4096.

При обучении наборы данных не изменялись, хотя в них не содержалась ответов на русском языке. Однако исследователи обнаружили [8] способность моноязыковых больших моделей обрабатывать многоязыковые задачи. Для сравнения двух моделей – *LLaVA* с энкодером *Saiga/Mistral 7b (LLaVA-Saiga/Mistral 7b)* и оригинальной *LLaVA* был применен следующий подход.

1 Перевести на русский язык три набора данных, представленных в таблице 1 (*VQAv2, GQA, POPE*).

2 Получить ответы модели *LLaVA-Saiga/Mistral 7b* на вопросы из входных наборов данных.

3 Перевести ответы модели на английский язык и сравнить с ответами оригинальной модели *LLaVA 1.5*.

Такой подход обусловлен отсутствием визуального вопросно-ответного набора данных для обучения на русском языке. В эксперименте было решено не использовать синтетические наборы данных во избежание потери способности обобщения, так как модель *Saiga/Mistral 7b* обучалась на синтетических данных. Оценки точности двух моделей представлены в таблице 3. Жирным шрифтом выделены оценки точности для модели *LLaVA* с текстовым энкодером *Saiga/Mistral 7b*.

Обучение длилось одну эпоху, как и предлагают авторы модели *LLaVA* [5], при этом был использован алгоритм низкоранговой адаптации *LoRA*. Шаг обучения, функции потерь и основные параметры алгоритмы *LoRA* при обучении не изменялись по сравнению с предложенными значениями авторами в оригинальной статье [5]. Во время проведения эксперимента, было отмечено, что весовые коэффициенты полученные с помощью алгоритмы *LoRA*, требуют около 80 Гб оперативной памяти графического процессора. Обучение проходило на одном графическом ускорителе с тензорными ядрами *NVIDIA A100*. Одна эпоха продлилась приблизительно одну неделю. Во время подготовки к процессу обучения была также замечена особенность модели *LLaVA*: обучить оригинальную модель удалось только на графическом ускорителе архитектуры *Ampere*, на других архитектурах графических ускорителей обучение оказалось невозможным из-за несовместимости рассматриваемой модели с архитектурой *CUDA*.

Таблица 3. Оценки точности оригинальной модели *LLaVA 1.5* и модели *LLaVA 1.5* с энкодером *Saiga/Mistral 7b*

Имя модели	Набор данных <i>VQAv2</i> , метрика <i>accuracy</i> [5]	Набор данных <i>GQA</i> , метрика <i>accuracy</i> [5]	Набор данных <i>POPE</i> , метрика <i>accuracy</i> [5]
<i>LLaVA 1.5</i>	0,8	0,63	0,85
<i>LLaVA-Saiga/Mistral</i>	0,72	0,53	0,79

Результаты, представленные в таблице 1, показывают, что изменение текстового энкодера в архитектуре на другом языке ухудшают точность модели. Однако, учитывая, что модель *LLaVA-Saiga/Mistral 7b* обучалась на наборе данных на другом языке, возможно, точность удастся улучшить на наборе реальных данных на русском языке.

Однако основные свойства и особенности модели сохранились, как было отмечено во время проведения ручного тестирования модели, вопросно-ответная система сохранила свою возможность вести многоходовые диалоги, отвечать на несколько вопросов заданных по одному изображению. Сохранилась возможность обобщения визуальных данных. Модель стала лучше различать кириллические рукописные символы на

изображениях, что позволяет более точно решать задачу распознавания символов на изображениях. Исходя из этих наблюдений, можно сделать вывод, что для улучшения точности модели и получения лучших значений показателя точности модели, чем у оригинальной модели *LLaVA*, требуется провести больше одной эпохи и подготовить набор данных для *BBOM* на русском языке. Возможно, его удастся синтезировать, аналогично набору данных *Saiga*, который использовался для обучения модели *Saiga/Mistral 7b*.

Заключение. Проведенный эксперимент показывает возможность обучения вопросно-ответной модели на основе архитектуры *LLaVA* методом замены энкодера. Учитывая результаты сравнения, можно сделать вывод, что современные большие языковые модели, способные решать задачи на языках, отсутствующих в обучаемой выборке, способны решать задачу векторного представления в вопросно-ответных моделях.

Список литературы

- [1] Демидова Л.А., Морозкин Н.А. Аспекты разработки архитектуры вопросно-ответной системы для обработки больших данных на основе нейросетевого моделирования // Вестник Рязанского государственного радиотехнического университета. 2023. No 86. С. 55–69.
- [2] Aishwarya Agrawal, Jiasen Lu, Stanislaw Antol, Margaret Mitchell, C. Lawrence Zitnick, Dhruv Batra, Devi Parikh VQA: Visual Question Answering // International Conference on Computer Vision (ICCV) 2015, 2015. pp. 1-13.
- [3] Husein Zolkepli, Aisyah Razak, Kamarul Adha, Ariff Nazhan Multi-Lingual Malaysian Embedding: Leveraging Large Language Models for Semantic Representations // arXiv.org, 2024, DOI:10.48550/arXiv.2402.03053
- [4] Sophie Xhonneux, David Dobre, Jian Tang, Gauthier Gidel, Dhanya Sridhar In-Context Learning Can Re-learn Forbidden Tasks // arXiv.org, 2024, DOI:10.48550/arXiv.2402.05723
- [5] Haotian Liu, Chunyuan Li, Qingyang Wu, Yong Jae Lee Visual Instruction Tuning // arXiv.org, 2023, DOI:10.48550/arXiv.2304.08485
- [6] Haotian Liu, Chunyuan Li, Qingyang Wu, Yong Jae Lee Improved Baselines with Visual Instruction Tuning // arXiv.org, 2023, DOI:10.48550/arXiv.2310.03744
- [7] Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Léo Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, William El Sayed Mistral 7B // arXiv.org, 2023, DOI:10.48550/arXiv.2310.06825
- [8] Mikhail Tikhomirov, Daniil Chernyshev Impact of Tokenization on LLaMa Russian Adaptation // arXiv.org, 2023, DOI:10.48550/arXiv.2312.02598

Авторский вклад

Демидова Лилия Анатольевна – руководство исследованием по оценке точности модели, постановка задачи исследования, описание общей постановки задачи вопросно-ответного моделирования

Морозкин Никита Андреевич – Реализация процесса обучения модели *LLaVA* с текстовым энкодером *Saiga/Mistral 7b*, сравнение рассматриваемой модели с оригинальной моделью *LLava*, анализ полученных результатов

TRAINING A QUESTION AND ANSWERING NEURAL NETWORK MODEL BASED ON LLaVA 1.5 MODEL ARCHITECTURE WITH SAIGA MISTRAL 7B ENCODER AND LOW-RANK ADAPTATION ALGORITHM LORA

L.A. Demidova

*Doctor of Technical Sciences,
Professor of the Department of
Corporate Information Systems,
Institute of Information
Technologies MIREA - Russian
Technological University,
Moscow, Russia;*

N.A. Moroshkin

*postgraduate student of the
Department of Corporate
Information Systems, Institute of
Information Technologies MIREA -
Russian Technological University,
Moscow, Russia;
orcid.org/0009-0002-8787-2452,*

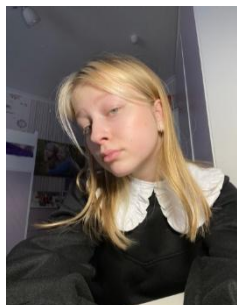
Annotation. In this work, a training algorithm for a visual question-answer neural network model is implemented based on the structure of the LLaVA 1.5 neural network model using the Mistral 7b text encoder, which allows improving the model's performance in the task of visual question-answer modeling and the LoRA low-rank adaptation algorithm, which speeds up the model learning process .

It is shown that the problem of question-answer modeling can be solved by neural network models using large language models, and a technique for accelerating the training of such models is described. The effectiveness of the algorithmic model is assessed and general aspects of training question-answer models on the most popular sets of question-answer data are shown.

Keywords: Visual question-answer modeling, large language models, low-rank adaptation LoRA, neural network model LLaVA 1.5, Saiga/Mistral 7b.

УДК 004.021:004.75

ИГРА ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ «ОСТРОВНОЕ ИССЛЕДОВАНИЕ: ГОЛОВОЛОМКИ И ТЕЛЕПОРТЫ»



А.Ю.Саливон
Учащийся Национального детского технопарка, учащаяся Высоковской средней школы им. Днепровской флотилии Пинского района



Е.Ю.Красникевич
Учащийся Национального детского технопарка, учащаяся гимназии №146 г. Минска



Г.О.Стешиц
Учащийся Национального детского технопарка, учащийся гимназии № 1 г. Солигорска



Ф.В. Усенко
инженер-программист кафедры инженерной психологии и эргономики БГУИР, магистр
f.usenko@bsuir.by



М.С. Ильясова
ассистент кафедры инженерной психологии и эргономики БГУИР, магистр
m.ilyasova@bsuir.by



Л.Р. Коркин
Заведующий лабораторией виртуальной и дополненной реальности Национального детского технопарка, магистр технических наук

А.Ю.Саливон

Обучается в Высоковской средней школы им. Днепровской флотилии Пинского района. Областью научных интересов является разработка приложений виртуальной и дополненной реальности.

Е.Ю.Красникевич

Обучается в гимназии №146 г. Минска. Областью научных интересов является разработка приложений виртуальной и дополненной реальности.

Г.О.Стешиц

Обучается в гимназии № 1 г. Солигорска. Областью научных интересов является разработка приложений виртуальной и дополненной реальности.

М.С. Ильясова

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов находится на пересечении информационных технологий и промышленной психологии.

Ф.В. Усенко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Областью научных интересов является моделирование пользователей информационных систем и разработкой интерфейсов.

Л.Р. Коркин

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с системами распознавания изображений.

Аннотация. Статья описывает разработку VR-игры в жанре квест-развлечения «Островное исследование: головоломки и телепорты». Авторы анализируют аналоги, описывают подробный сценарий игры, а также процесс разработки, включая написание программного кода. Также обсуждается тестирование приложения, гарантирующее его надежную работу, и целевая аудитория.

Ключевые слова: квест, VR, Unity, головоломки, Blender 3D, скрипт

Введение. Виртуальная реальность представляет собой технологию, создающую компьютерную среду, в которую пользователь может погружаться, полностью отделяясь от реального мира. Перенос в виртуальную реальность обеспечивается различными устройствами, включая VR-очки, которые обеспечивают визуальное погружение в виртуальный мир, а также разнообразные датчики движения и контроллеры для взаимодействия в этом пространстве.

Для реализации проекта был выбран жанр квест-развлечение. Он предусматривает решение головоломок, загадок и задач, требует умственных усилий от игрока. Задания могут быть текстовыми, как интерактивная книга, и графическими, связанными с поиском объектов на картинке. Игры в этом жанре, в первую очередь, занимают голову, а не руки.

Особенности квестов – ролевая игра, цель в которой – решить головоломки, выполнить задания и дойти до финала игры.

Привлекательность квестов, прежде всего в том, что на первый план в играх такого жанра выходит не качество графики или реалистичность происходящего на мониторе, а сюжет, характер персонажа, которым управляет пользователь.

Игрок участвует в тактических поединках и доходит до конца игрового сюжета. Рост возможностей персонажа полностью контролируется геймером и зависит только от него.

Для этого был проведен анализ схожих игр.

Moss – динамичное приключение с элементами головоломки для одного игрока, созданное специально для VR. Она сочетает классические элементы игры – интересных персонажей, захватывающую боевую систему, исследование. В *Moss* вы познакомитесь с маленькой мышкой, мечтающей об огромном мире. Вместе вы будете путешествовать по забытым царствам, решать хитрые загадки и сражаться с опасными противниками. Но наша игра, в отличие от *Moss*, который разработан на игровом движке *Unreal Engine 4*, использует движок *Unity*, который является более простым и удобным в использовании и имеет огромную библиотеку ассетов и плагинов. Также *Moss* не поддерживает русский язык [1].

Return to Monkey Island – компьютерная игра в жанре квест, является шестой игрой в серии *Monkey Island*. Игроки отправятся в путешествие вместе с Гаябрушем Трипвудом, решая головоломки, встречая интересных персонажей, исследуя тропические острова и разгадывая загадки [2].

Myst – компьютерная игра в жанре графического квеста. Игровой процесс *Myst* представляет собой путешествие по интерактивному миру от первого лица. Игроки могут взаимодействовать с определёнными объектами на экране, щёлкая по ним или перетаскивая их. В этой игре игрок попадает на таинственный остров, решает головоломки и раскрывают предысторию персонажей [3].

Цели и задачи. Создать игру в жанре квест-развлечение «Островное исследование: головоломки и телепорты» в VR среде.

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 сцен;
- 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# [4].

Разработка сценария игры. Для реализации игры был выбран такой жанр, как головоломки, специализирующийся на развитии умственных способностей и поддержании здорового ума человека, совершенствовании и тренировки мышления.

В начале игры пользователь оказывается в лаборатории, в которой ведётся работа по изучению телепортации. Задача игрока на данном этапе – разблокировать портал и телепортироваться. Переместившись в пещеру на необитаемом острове, игроку посредством решения логических головоломок необходимо попасть к особняку, расположенному на другом берегу острова. Исследуя локации, игрок будет находить подсказки, благодаря которым в дальнейшем он сможет найти местонахождение обратного портала, то есть пройдёт игру.

По мере прохождения игры игрок будет сталкиваться с различными головоломками, часть из которых представлена ниже.

Головоломки в лаборатории:

- 1 На столе собрать пазл в форме портала из 12 фрагментов.
- 2 Используя ультрафиолетовую лампу, прочитать появившийся под ней текст на пазле, который соответствует подсказке о включении портала в розетку.
- 3 Разблокировать портал, выбрав на экране 5 нужных символов из предложенных (символы будут находиться на разных предметах в комнате).

Головоломки на острове:

1 В пещере на стенах размещены рисунки, связанные с костром. Надо раскопать угли в кострище и найти карту острова, на которой крестиком отмечен особняк, а кружочком лодка.

2 Найти по карте пирс с лодкой, в которой находится ящик, закрытый на кодовый замок. Необходимо ввести код, соответствующий расположению брёвен в пещере.

3 Следуя по карте, обойти остров к особняку. Используя молоток, полученный из ящика ранее, извлечь гвозди из досок, для получения доступа на территорию особняка.

4 Исследовать кухню и ванну. В морозильной камере на кухне лежит лёд, который надо растопить с помощью фена, находящегося в ванной. После чего можно будет взять ключ от тайной комнаты №1.

5 Открыть тайную комнату №2, выдвинув из стеллажа правильную книгу. Книжный шкаф отъедет и появится библиотека с роялем. Задача игрока – сыграть на пианино мелодию, которая играет в комнате (взаимодействие с пианино осуществляется посредством 6 клавиш).

6 В кабинете надо разблокировать ноутбук, и тогда запустится принтер и напечатается подсказка о том, как попасть в тайную комнату №3.

7 Пройти лабиринт на чердаке для открытия выхода к домику на дереве. После чего совместить шестерёнки в правильной последовательности для открытия двери.

8 Заметить во время прохождения лабиринта на стене картину-карту, которая подскажет, где закопан ключ. Для выявления месторасположения ключа от комнаты с порталом, необходимо исследовать лабиринт на наличие подсказок и откопать лопатой, находившейся на стене домика, ключ.

9 Необходимо разблокировать рубильник, который будет висеть на стене, для этого необходимо правильно ответив на 10 вопросов о географии, правилах безопасности и поведения на открытой местности, после чего, запустится портал.

Для определения расположения объектов в особняке был разработан план дома (рис.1). После чего были созданы 3д модели этих предметов с помощью среды моделирования *Blender* [5].

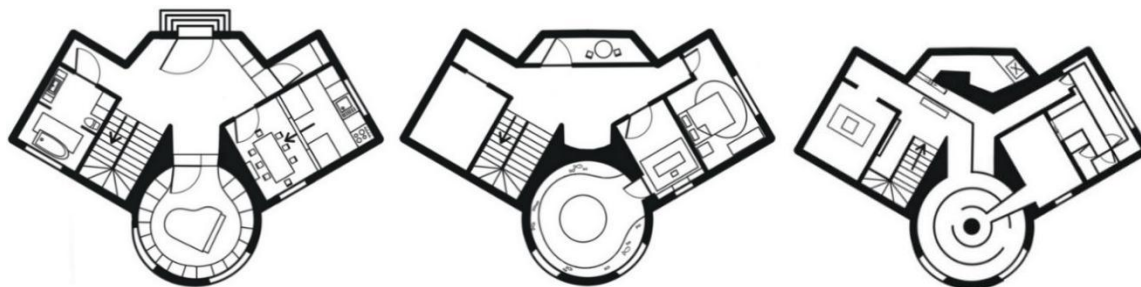


Рисунок 1. Художественное изображение особняка (вид сверху)

Разработка программного кода. Для того чтобы свободно ходить по карте, перемещение игрока в игре должно происходить при взаимодействии с джойстиком на контроллере. Для реализации этой механики был написан программный код, благодаря которому игрок может свободно передвигаться по карте, а также позволяет внедрить гравитацию (рис. 2) [6].

```
public class JoystickLocomotion : MonoBehaviour
{
    public SteamVR_Input_Sources handType;
    public SteamVR_Action_Vector2 joystickAction;

    public float speed = 3.0f;

    private CharacterController characterController;
    private const float gravityForce = 9.81f;
    private float verticalVelocity = 0f;

    void Start()
    {
        characterController = GetComponent<CharacterController>();
    }

    void Update()
    {
        if (!characterController.isGrounded)
        {
            verticalVelocity -= gravityForce * Time.deltaTime;
        }
        else
        {
            verticalVelocity = 0;
        }

        Vector2 joystickValue = joystickAction.GetAxis(handType);
        Vector3 moveDirection = new Vector3(joystickValue.x * speed, verticalVelocity, joystickValue.y * speed);
        moveDirection = Quaternion.Euler(0, Camera.main.transform.eulerAngles.y, 0) * moveDirection;
        characterController.Move(moveDirection * Time.deltaTime);
    }
}
```

Рисунок 2. Код для перемещения игрока

Были проведены модульное, интеграционное и системное тестирования для обеспечения надёжной и стабильной работы приложения. Все найденные в ходе тестирования ошибки и недочёты были выявлены и исправлены, а программный продукт готов к использованию.

Заключение. В результате разработки все поставленные цели были достигнуты, задачи были выполнены. В итоге получилось реализовать квест-игру в VR «островное исследование: головоломки и телепорты». Этот продукт может быть нацелен на любую возрастную категорию, но в основном подойдет для людей в возрасте от 8+ лет, также его может использовать человек любого вероисповедания и национальности.

Список литературы

- [1] Moss [Электронный ресурс]. Режим доступа: <https://elonium.by/product/moss/>. Дата доступа: 28.01.2024.
- [2] Return to Monkey Island [Электронный ресурс]. Режим доступа: https://ru.wikipedia.org/wiki/Return_to_Monkey_Island. Дата доступа: 28.01.2024.
- [3] Myst [Электронный ресурс]. Режим доступа: <https://elonium.by/product/myst/>. Дата доступа: 28.01.2024.
- [4] Unity Documentation [Электронный ресурс]. Режим доступа: <https://docs.unity.com/>. Дата доступа: 28.01.2024.
- [5] Blender [Электронный ресурс]. Режим доступа: <https://www.blender.org/>. Дата доступа: 28.01.2024.
- [6] C# Tutorial [Электронный ресурс]. Режим доступа: <https://www.w3schools.com/cs/index.php>. Дата доступа: 30.01.2024.

Авторский вклад

Авторы внесли равноценный вклад.

DEVELOPMENT OF A VIRTUAL TRUCK SIMULATOR

A.Y.Salivon

Student of the National Children's Technopark, a student of the Vysokovskaya Secondary School named after The Dnieper flotilla of the Pinsky district

E.Y.Krasnikievich

Student of the National Children's Technopark, a student of gymnasium No.146 in Minsk

G.O.Steshits

Student of the National Children's Technopark, a student of gymnasium No. 1 in Soligorsk

P.V. Usenko

Software engineer, Department of Engineering Psychology and Ergonomics BSUIR, M.Sc.

M.S. Ilyasova

*Assistant of the Department of Engineering Psychology and Ergonomics of BSUIR, M.Sc.
m.iliasova@bsuir.by*

L.R. Korkin

Head of the Virtual and Augmented Reality Laboratory at the National Children's Technopark, Master of Technical Sciences

Abstract. The article describes the development of a VR game in the genre of quest entertainment «Island Exploration: puzzles and teleports». The authors analyze analogs, describe a detailed scenario of the game, as well as the development process, including writing program code. Testing of the application, which guarantees its reliable operation, and the target audience are also discussed.

Keywords: quest, VR, Unity, puzzles, Blender 3D, script

УДК 65.012.123

ОДИН ИЗ АЛГОРИТМОВ ОЦЕНКИ СТИЛЯ УПРАВЛЕНИЯ ТРАНСПОРТНЫМ СРЕДСТВОМ



Д.Д. Лифарева
Студентка кафедры
радиофотонных и
микроволновых технологий
Казанского национального
исследовательского
технического университета
им. А.Н. Туполева
diana.Lifareva@mail.ru



Е.А. Курьянова
Студентка кафедры
радиофотонных и
микроволновых технологий
Казанского национального
исследовательского
технического университета
им. А.Н. Туполева
katerinka47328@yandex.ru



Д.А. Веденькин
Доцент кафедры
радиофотонных и
микроволновых технологий
КНИТУ-КАИ им. Туполева,
кандидат технических наук,
доцент
denis_ved@mail.ru

Д.А. Веденькин

Окончил Казанский государственный технический университет им. А.Н. Туполева в 2005 году по направлению Радиотехника. Канд. техн. наук, доцент. Область научных интересов связана с антеннами, СВЧ устройствами и их технологиями, телекоммуникационными технологиями, радиотехникой.

Д.Д. Лифарева

Обучается в Казанском национальном исследовательском техническом университете им. А.Н. Туполева-КАИ по направлению Радиотехника. Профиль: Микроволновые технологии. Область научных интересов связана с радиотехническим оборудованием, а также изучением микроволновых процессов и СВЧ-приборов.

Е.А. Курьянова

Обучается в Казанском национальном исследовательском техническом университете им. А.Н. Туполева-КАИ по направлению Радиотехника. Профиль: Микроволновые технологии. Область научных интересов связана с радиотехническим оборудованием, а также изучением микроволновых процессов и СВЧ-приборов.

Аннотация. Выполнено создание алгоритма оценки стиля вождения автомобиля, обеспечивающего качественный анализ состояния водителя и качества его подготовки, алгоритм реализован на базе среды разработки *LabView*.

А также реализована система сбора анализируемых в дальнейшем данных с автомобиля и его окружения (показатели дальности до объектов, погодные условия, считывание дорожных знаков и т.д.).

Ключевые слова: *LabView*, оценка стиля вождения автомобиля, сбор данных, анализ получаемых данных.

Введение. Стиль вождения – это типичные, устоявшиеся особенности дорожного поведения водителя, система способов и средств осуществления им деятельности, а также типичные для водителя эмоциональные переживания и взаимодействия с другими участниками дорожного движения, предопределяющие уровень безопасности (аварийности) водителя и определяемые рядом факторов. Существует множество систем оценки стиля вождения, используемых в различных отраслях, таких как логистика,

страхование, автомобильная промышленность и другие. Кроме того, оценка стиля вождения может быть использована для разработки программ обучения для водителей.

Современные автомобили представляют собой комплексные высокотехнологичные системы, оборудованные множеством датчиков и различных управляющих систем, которые включают в себя использование внешних информационных данных, таких как географические координаты от спутниковой навигации, с последующей ассоциацией с картографическими данными. Кроме того, такие системы включают возможность считывания информации о дорожных знаках, разметке, оценке бокового интервала, дистанции до других объектов, а также скорости приближения к другим участникам дорожного движения. Анализ стиля вождения играет ключевую роль в оценке качества подготовки водителей, поскольку является важным и объективным показателем. Предоставление такой оценки позволяет организациям оптимизировать использование автотранспорта, сократить расходы на обслуживание и снизить вероятность возникновения дорожных происшествий.

Предлагаемый подход предусматривает возможность проведения оценки стиля вождения в различных ситуациях, таких как трассы, городские улицы и жилые районы. Для каждой из этих ситуаций предусмотрены определённые допустимые пределы изменения параметров. Кроме того, возможно рассчитать общую оценку стиля вождения путём комплексной оценки всех параметров. Всё это может быть реализовано с использованием оптических [1], микроволновых [2] технологий с учётом требований ЭМС, подходы к которым описаны в [3].

Сбор данных. Были выделены следующие критические данные для сбора информации с автомобиля:

1 Скорость движения и разница с разрешённой скоростью. Позволяют оценить, насколько водитель соблюдает ограничения скорости, а также выявить возможные нарушения.

2 Показатели, связанные с двигателем: обороты двигателя, массовый расход воздуха, могут указывать на агрессивный стиль вождения.

3 Интенсивность ускорения и торможения может служить показателем динамического поведения водителя. Резкое ускорение и торможение могут указывать на агрессивный и нестабильный стиль вождения.

4 Отклонение рулевого колеса, связанное с различными скоростями движения, позволяет оценить стабильность руления и контроль за автомобилем.

5 Расстояние до препятствий и боковые интервалы отражают уровень внимательности водителя на дороге.

6 Навигационные данные, погодные условия и состояние дорожного покрытия могут влиять на стиль вождения. Например, вождение в плохих погодных условиях или на гладком дорожном покрытии может требовать изменения стиля вождения.

На рисунке 1 представлена упрощённая структурная схема устройства для сбора данных.

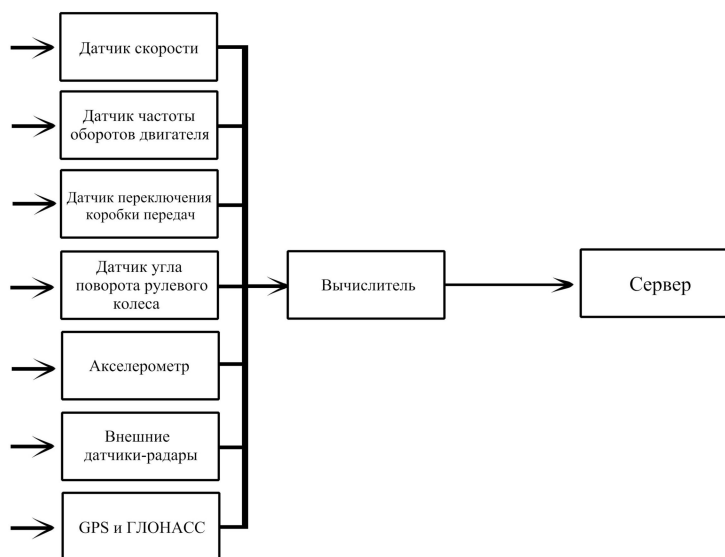


Рисунок 1. Структурная схема устройства для сбора данных

Частота опроса датчиков и сбора данных зависит от объёма памяти и требований по оценке скорости изменения параметров. Типовые значения периодичности опроса составляют около 2-5 секунд, но могут быть настроены в соответствии с конкретными потребностями и требованиями системы оценки стиля вождения.

Важно отметить, что для анализа рекомендуется использовать комплексный подход, учитывая не только отдельные данные, но и их взаимосвязь, а также контекстуальные условия ситуации на дороге.

Анализ данных.

LabVIEW – это среда разработки и платформа для выполнения программ, созданных на графическом языке программирования «G» фирмы *National Instruments*.

Для корректного анализа данных были выявлены следующие необходимые этапы:

1 Выявление ситуации, в которой находится транспортное средство в момент считывания данных. Определение погодных условий, сложности трассы и состояния дорожного полотна.

2 Определение верхней и нижней допустимых границ для каждого параметра. Например для скорости движения верхней границей будет являться скорость движения скорой помощи в черте города, а нижней – скорость ученической машины.

3 Получение данных с прибора.

4 Сравнение полученных данных с верхней и нижней границами допустимых значений с учётом весовых коэффициентов, полученных при анализе погодных условий.

5 Вывод итоговой оценки стиля вождения.

Рассмотрим алгоритм оценки для значения средней скорости движения. Наглядно алгоритм вычисления оценки представлен на рисунке 2.

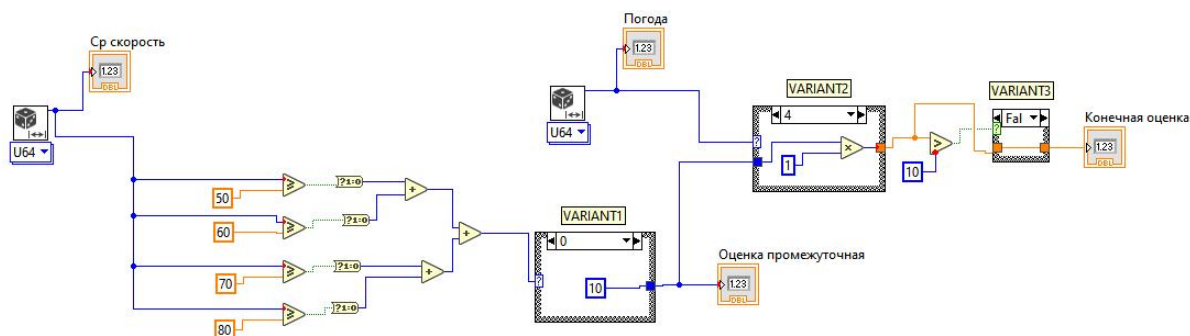


Рисунок 2. Модель расчёта оценочного значения для оценки средней скорости автомобиля

Значения скорости сравниваются с пороговыми значениями, соответствующими определённым оценочным значениям. Так, например если скорость равна 65 км/ч, значение превышает 50 км/ч и 60 км/ч, а значит на выходе двух блоков сравнения присутствуют единицы. Далее результаты сравнения суммируются. В нашем примере результат равен 2. Полученное число поступает в блок *VARIANT1*, в котором каждому возможному полученному значению присваивается определённое оценочное значение. На индикатор выводится промежуточная оценка.

Далее происходит вычисление конечной оценки с учётом весовых коэффициентов, происходит это следующим образом: промежуточная оценка поступает в блок *VARIANT2*, где в зависимости от погодных условий умножается либо делится на весовой коэффициент. Весовые коэффициенты зависят от полученных данных о погоде, сложности трассы и состоянии дорожного полотна. В дальнейшем результат вычисления сравнивается в блоке *VARIANT3* с максимально допустимым значением оценки и если результат превышает его, ему присваивается значение 10, если результат меньше максимально допустимого, то он остаётся фактическим.

Алгоритм аналогичен и для других параметров оценки стиля вождения, таких как: обороты двигателя, расстояние до препятствий, интенсивность ускорения и торможения и т.д. С той лишь разницей, что верхние и нижние границы у них будут другими. Верхние и нижние границы ещё двух параметров приведены ниже:

1 Обороты двигателя: верхняя граница – 4500 об/мин., нижняя граница – 2000 об/мин.;

2 Расстояние до препятствий: верхняя граница – 2 м., нижняя граница – 0.8 м.

Также следует отметить, что рассмотренный выше пример включает оценку скорости в пределах города. Для загородной зоны границы оценки будут иными.

Когда оценки всех параметров выведены вычисляется среднее арифметическое для всех оценок параметров. Это и есть конечный результат оценки стиля вождения автомобиля.

Заключение. Итогом работы является создание устройства, осуществляющего сбор необходимой для анализа информации с датчиков автомобиля и систем *GPS* и *ГЛОНАСС*. А также разработан алгоритм для вычисления оценки стиля вождения на базе программной среды *LabView*.

Список литературы

[1] Лустина, А.А. Векторный измеритель доплеровского сдвига частоты локационного сигнала на основе тандемной амплитудно-фазовой модуляции и волоконной брэгговской решетки, поддерживающей поляризацию / А.А. Лустина, П.Е. Денисенко, Р.М. Шагвалиев, А.А. Иванов, О.Г. Морозов и др. // Электроника, фотоника и киберфизические системы. 2023. Т.3, №2. С. 32-43

[2] Веденькин, Д.А. Антенные решетки, сфокусированные по широкополосному сигналу /Д.А. Веденькин, Ю.Е. Седельников // Физика волновых процессов и радиотехнические системы. 2015. Т.18. №3. С. 23-30.

[3] Веденькин, Д.А. Оценка электромагнитной совместимости радиотехнического оборудования перспективных беспилотных летательных аппаратов на этапах разработки /Д.А. Веденькин, Ю.Е. Седельников, В.Е. Латышев // Вестник Поволжского государственного технологического университета. Серия: Радиотехнические и инфокоммуникационные системы. 2014. №5(24). С. 57-64.

Авторский вклад

Веденькин Денис Андреевич – руководство исследованием по оценке качества стиля вождения. Научный руководитель.

Лифарева Диана Дмитриевна – разработка алгоритма оценки стиля вождения автомобиля, реализованного на базе среды разработки LabView.

Курьянова Екатерина Антоновна – реализация системы сбора анализируемых в данных с автомобиля и его окружения.

ONE OF THE ALGORITHMS FOR ASSESSING VEHICLE DRIVING STYLE

D.D.Lifareva

*Student of the Department of
Radiophotonic and Microwave of
KNRTU
n.a. A.N. Tupolev-KAI*

E.A. Kuryanova

*Student of the Department of
Radiophotonic and Microwave of
KNRTU
n.a. A.N. Tupolev-KAI*

Vedenkin D.A

*Associate Professor of the
Department of Radiophotonic and
Microwave of KNRTU
n.a. A.N. Tupolev-KAI, PhD of
Technical Sciences, Associate
Professor*

Abstract. An algorithm for assessing the driving style of a car was created, providing a qualitative analysis of the driver's condition and the quality of his training; the algorithm was implemented based on the LabView development environment.

A system has been implemented for collecting further analyzed data from the car and its surroundings (distance indicators to objects, weather conditions, reading road signs, etc.).

Keywords: LabView, car driving style assessment, data collection, analysis of received data.

УДК 004.891

АЛГОРИТМЫ ПОВЫШЕНИЯ ЧИТАБЕЛЬНОСТИ ТЕКСТА НА ИЗОБРАЖЕНИЯХ НИЗКОГО КАЧЕСТВА С ПРИМЕНЕНИЕМ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ



Сачивко Н.С.
Студент факультета-
компьютерных систем и сетей
БГУИР
skriercoder@gmail.com



Калугина М.А.
Доцент кафедры
информатики БГУИР,
кандидат физико-
математических наук,
доцент
marina_kalugina@list.ru

Н.С. Сачивко

В 2020 году окончил Браславскую гимназию. Неоднократно становился победителем олимпиад по математике, информатике, финансовой грамотности, английскому языку. Интересуется информационной безопасностью, информатикой и математикой.

М.А. Калугина

Окончила Белорусский государственный университет. Область научных интересов связана с исследованием проблем метрической теории диофантовых приближений зависимых величин и приложений математических методов к нейросетевому анализу.

Аннотация. В рамках данной статьи рассмотрена проблема нечитаемого текста на изображениях низкого качества. Предложено два алгоритма повышения его читабельности: алгоритм увеличения размера изображения и алгоритм снижения степени его размытости. Описаны устройства моделей используемых нейросетей, приведены результаты работы предложенных алгоритмов.

Ключевые слова: машинное обучение, глубокое обучение, нейросети, нейросетевые модели, распознавание текста, увеличение изображений, повышение качества изображений, устранение размытия, повышение читабельности текста.

Введение. Не так давно фотография являлась непозволительной роскошью для многих людей, сам процесс съёмки занимал продолжительный период времени, а качество таких фотографий было достаточно низким. Затем появились фотоаппараты, использующие фотоплёнку; они уже были доступны большому количеству людей, но требовалось немало усилий и времени для получения готовых фотографий. Позже появились первые фотокамеры моментальной печати «Polaroid» с черно-белыми снимками.

Сегодня, пожалуй, каждый человек хоть раз в жизни сталкивался с цифровой фотографией в том или ином виде: полученной с помощью полноценного цифрового фотоаппарата, встроенной в иное устройство фотокамеры или даже сканера. Такой метод получения изображений и неразрывно связанных с ними видеозаписей встречается повсеместно: для любительской фотографии с использованием смартфонов, в криминалистической фотографии, для автоматической фиксации нарушений правил

дорожного движения, в целях оцифровки медицинских карт, для регистрации происшествий на дорогах (например, с помощью автомобильных видеорегистраторов). Безусловно, часто такие изображения (к слову, видеозаписи можно рассматривать как серию изображений) содержат важные сведения именно в текстовом виде, будь то номера автомобилей, рукописные медицинские сведения или важные документы.

К сожалению, не бывает идеальных систем фото- или видеофиксации, изображения не могут содержать в себе неограниченно детализированные сведения, а некоторые камеры способны производить лишь «замыленные» изображения низкого разрешения. Другими словами, всё в нашем мире конечно, в том числе и возможности цифровых фото- и видеокамер. Нередко текст на изображениях оказывается нечитаемым за счёт низкого разрешения последнего. Также не исключена ситуация, когда изображение получается размытым по той или иной причине. Если такое изображение попадает в систему, которая распознаёт текст в автоматическом режиме, то из-за низкого качества изображения возникает необходимость в его проверке человеком в ручном режиме. Впрочем, бывают случаи, когда не каждый человек способен распознать текст на таких изображениях – приходится либо заново производить фото- или видеосъёмку, если такая возможность имеется, либо просто пропускать изображение, что часто приводит к определённым потерям, а иногда и к критическим последствиям.

Один из подходов к решению проблемы распознавания текста на изображениях низкого качества – повышение качества изображения. Например, можно увеличить его разрешение, снизить степень размытости, устранить шум. Однако возникает проблема, связанная со сложностью данных операций: на самом деле, задача восстановления исходного изображения из размытого или из изображения с низким разрешением имеет слишком большое количество решений [1], то есть получить точно исходное изображение из «испорченного» на практике невозможно. Именно поэтому предлагается использовать нейросети и методы глубокого обучения для решения указанных задач.

Описание алгоритмов и обзор используемых технологий. Предполагается, что исходное изображение представлено в формате *RGB* и имеет размер по крайней мере 30×30 пикселей. В статье рассматривается два алгоритма: увеличения размера изображения и уменьшения степени его размытости.

Для помощи в решении поставленных задач взят язык программирования *Python*, библиотеки *Pillow*, *PyTorch* 2.1, *Torchvision* 0.16 и *Matplotlib*. В качестве аппаратного обеспечения используется среднестатистический домашний компьютер.

Важной особенностью библиотеки *Torchvision* является способ обработки изображений: встроенная функция *to_tensor* преобразует изображение в тензор, при этом стандартный диапазон яркости цветов от 0 до 255 приводится к диапазону $[0.0; 1.0]$. Этим и обусловлен выбор последнего слоя нейронных сетей: наиболее подходящим является слой активаций *Sigmoid*, ведь именно он обеспечивает корректный диапазон значений.

Функция активации *Sigmoid* определяется следующим образом:

$$s(x) = \frac{1}{1 + e^{-x}}.$$

Также будет применяться функция активации *ReLU*:

$$r(x) = \max(0, x).$$

Описание модели для увеличения размера изображения. Первая задача, которую необходимо решить, – это увеличение размера изображения. Для получения результата было принято решение использовать комбинацию свёрточной нейронной сети

(*Convolutional Neural Network* или *CNN*) и такой же сети, работающей в обратном процессе (*Deconvolutional Neural Network* или *DNN*).

Поскольку исходное изображение представлено в формате *RGB*, количество каналов на входе и на выходе нейросети принимается равным трём.

В ходе тестирования было установлено, что модель со следующими слоями является наиболее оптимальной (стрелкой указывается переход от исходного количества каналов к количеству каналов на выходе слоя):

1 *Deconvolution*: $3 \rightarrow 75$, ядро 6×6 , шаг 2.

2 *ReLU*.

3 *Convolution*: $75 \rightarrow 75$, ядро 4×4 , шаг 1.

4 *ReLU*.

5 *Convolution*: $75 \rightarrow 75$, ядро 4×4 , шаг 1.

6 *ReLU*.

7 *Deconvolution*: $75 \rightarrow 3$, ядро 6×6 , шаг 2.

8 *Sigmoid*.

Такая модель обеспечивает увеличение размера изображения в 4 раза.

Стоит отметить, что размеры ядра свёртки и шага для слоёв «деконволюции» (то есть слоёв, которые работают в обратном процессе по отношению к слоям свёртки, англ. «*deconvolution*») выбраны неслучайно: всё для того, чтобы избежать эффекта «шахматной доски». Если бы размер ядра свёртки не делился без остатка на шаг «деконволюции», то в результате этого процесса мог бы возникнуть «неравномерный нахлест» (англ. «*uneven overlap*») фрагментов, который и привёл бы к возникновению указанного эффекта [2]. Такой эффект заметен невооружённым глазом, как представлено на рисунке 1.

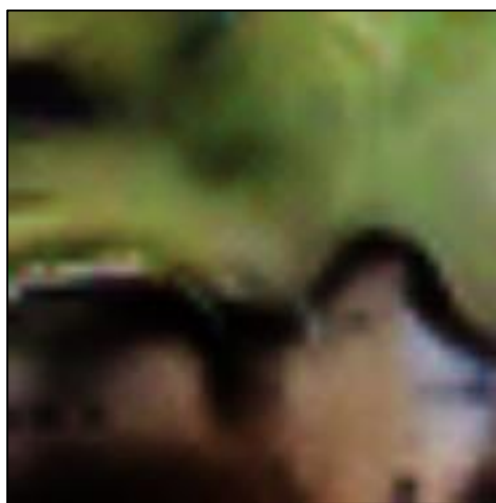


Рисунок 1. Пример эффекта «шахматной доски» [3]

Выбор набора данных для обучения модели для увеличения размера изображения. Изначально планировалось использовать набор изображений, содержащих части страниц общедоступных документов [4] – пример такого изображения представлен на рисунке 2.

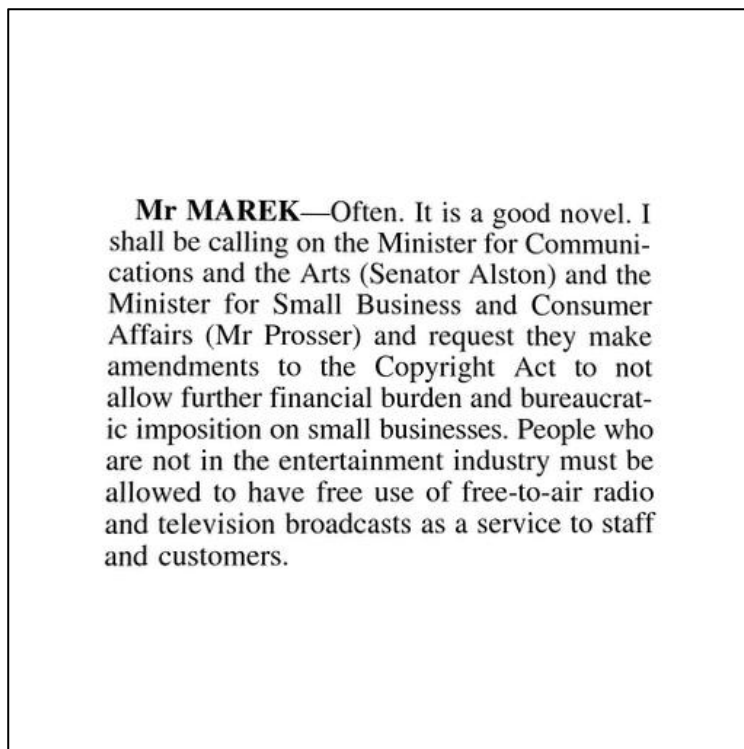


Рисунок 2. Пример исходного изображения (BMVC)

Такой набор данных, безусловно, хорош в случае, когда модель должна обрабатывать только документы. Но у такой модели есть один существенный недостаток: она преобразует все изображения, в том числе цветные, в чёрно-белые. По этой причине было решено использовать другой набор данных.

Например, существует набор синтетически сгенерированных символов [5] – они также являются чёрно-белыми, но все изображения в наборе содержат всего два цвета: либо полностью белый, либо полностью чёрный, что позволяет перекрасить их. Пример исходного символа из данного набор представлен на рисунке 3.

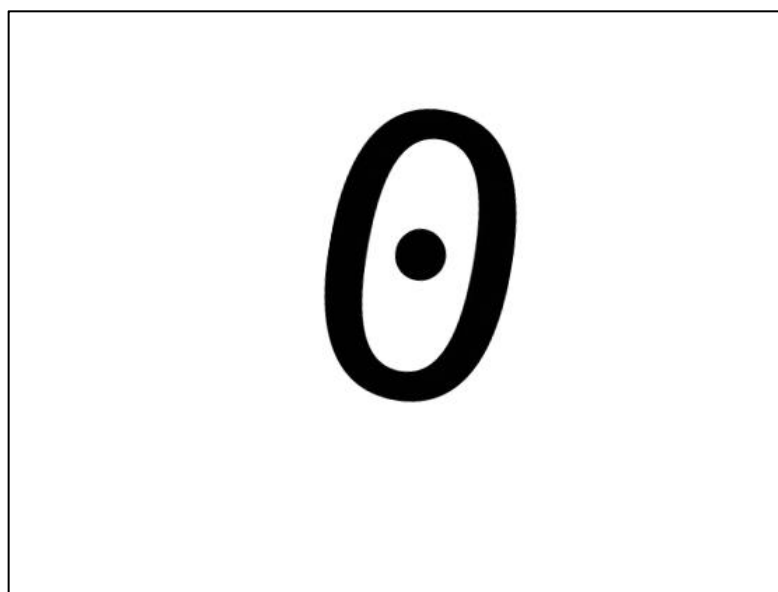


Рисунок 3. Пример исходного изображения (Synthetic Character Set)

С указанным набором данных были проделаны следующие операции:

- 1 Все изображения обрезаны так, чтобы по краям не оставалось белого фона.
- 2 Чёрный цвет в 20% случаев заменён на случайный.
- 3 Полученные изображения объединены в коллекции, причём в 20% случаев белый фон заменён случайным цветом, а в 50% случаев итоговое изображение повернуто на случайный угол.

4 Итоговые коллекции обрезаны до одинакового размера.

Примеры изображений из полученного набора (назовём его «исходный набор») представлены на рисунке 4.

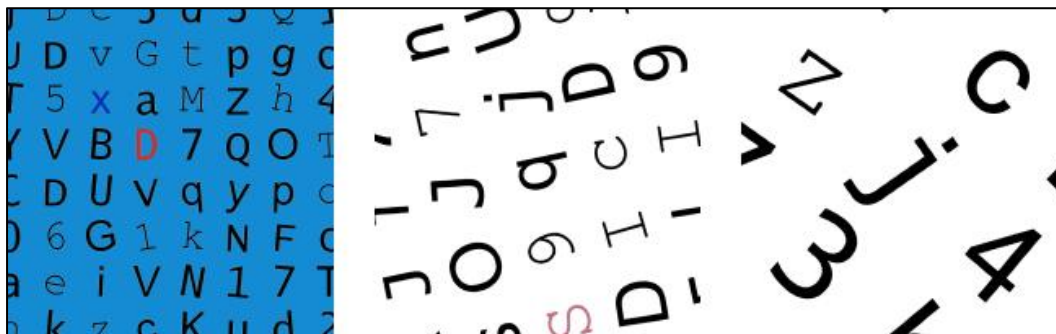


Рисунок 4. Пример изображений исходного набора

Изображения из исходного набора были уменьшены в 4 раза с использованием следующих методов:

- *BOX*;
- *HAMMING*;
- *LANCZOS*;
- *BICUBIC*;
- *BILINEAR*.

Примеры изображений из полученного набора (назовём его «итоговый набор») представлены на рисунке 5.

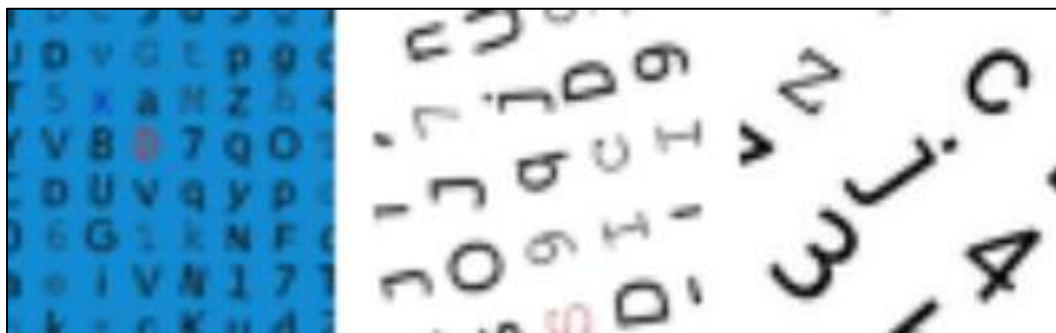


Рисунок 5. Пример изображений итогового набора

Для обучения созданной модели изображения из исходного и итогового наборов были собраны в пары, 70% таких пар отобраны случайным образом непосредственно для обучения модели, остальные 30% используются для валидации результатов.

Стоит отметить, что данные не были нормализованы. Хотя процесс нормализации и позволял достигнуть меньшей ошибки модели, но это также зачастую приводило к искажению цветов изображения в процессе обработки, что в итоге негативно сказывалось на восприятии человеком результата.

Обучение модели для увеличения размера изображения и достигнутые результаты. Модель обучалась из начального состояния, которое определялось случайным образом с помощью метода инициализации Ксавье (*Xavier* или *Glorot initialization*) [6].

Для обучения использовался модифицированный алгоритм стохастического градиентного спуска, именованный *Adam*. В ходе тестирования установлено, что именно данный алгоритм обеспечивает хорошую сходимость. В качестве меры ошибки была выбрана среднеквадратичная ошибка (*MSELoss*), которая определяется следующим образом:

$$l(X, Y) = \frac{\sum_{i=1}^N (X_i - Y_i)^2}{N},$$

где $N = |X| = |Y|$; X – исходный набор данных, Y – итоговый набор данных.

Модель обучалась в течение 600 эпох (одна эпоха означает, что все данные из созданного набора были использованы для обучения ровно один раз). График зависимости среднеквадратичной ошибки (оранжевым – ошибка на данных для валидации, синим – на тренировочных данных) от эпохи представлен на рисунке 6.

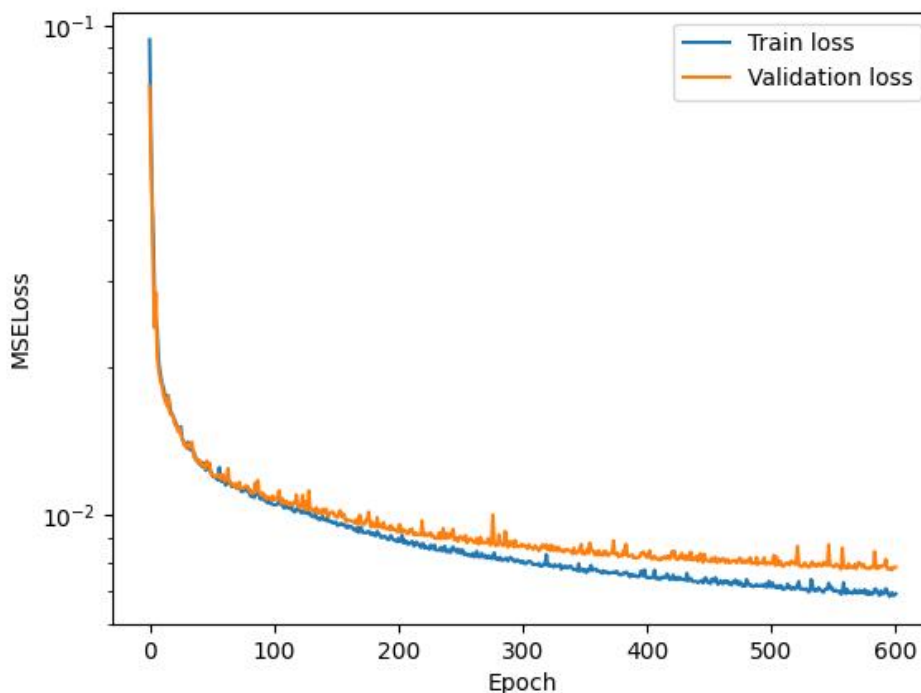


Рисунок 6. Ошибка на валидационных (оранжевым) и тренировочных (синим) данных, логарифмическая шкала

Пример обработки уменьшенных изображений представлен на рисунке 7, более детальное сравнение – на рисунке 8.

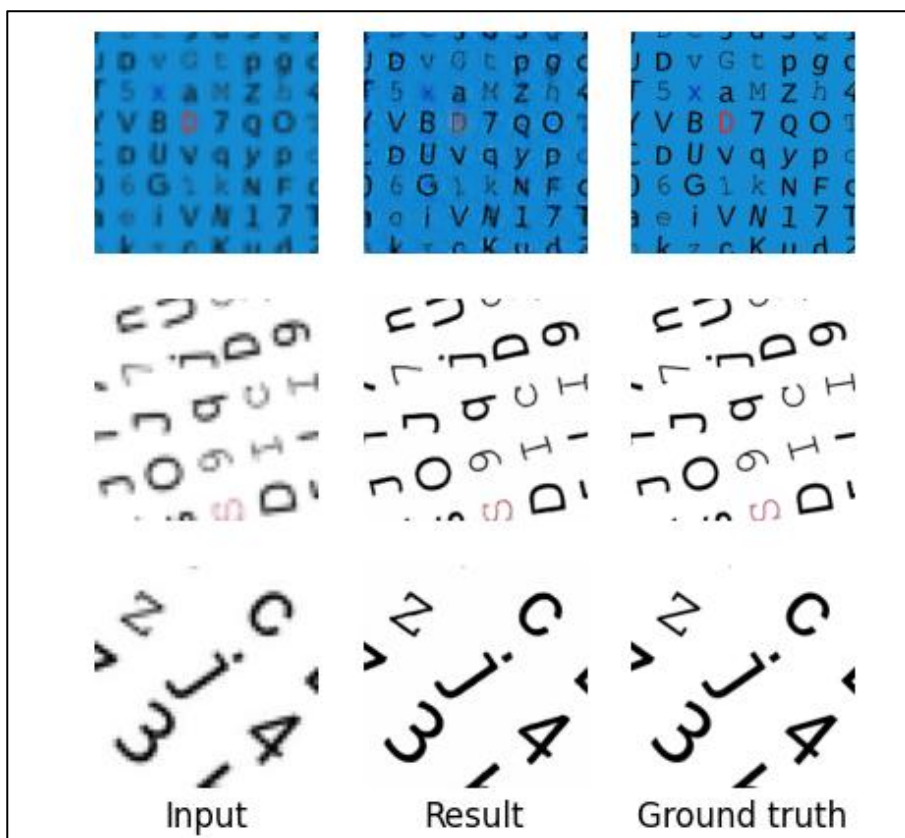


Рисунок 7. Пример обработки уменьшенных изображений

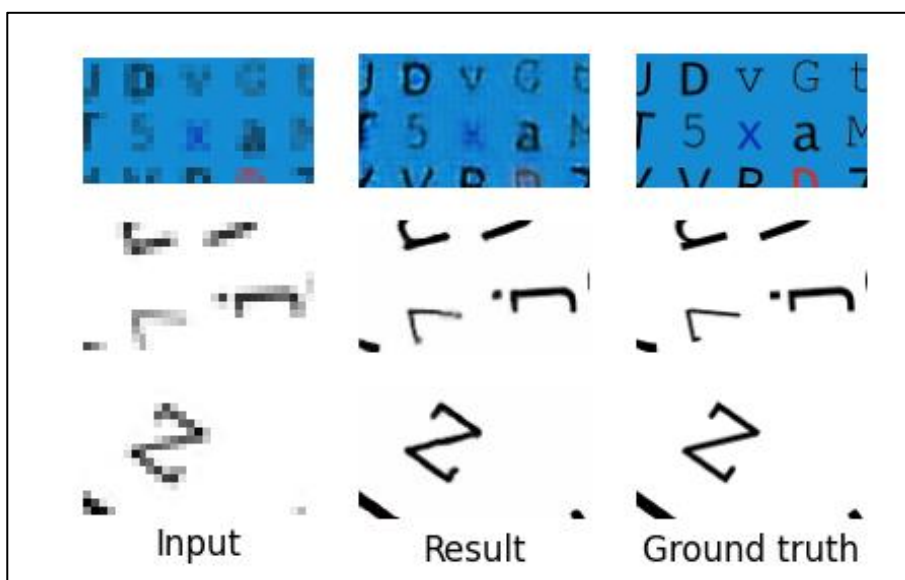


Рисунок 8. Детальное сравнение результатов обработки уменьшенных изображений

Видно, что модель корректно выполняет свою функцию: текст становится более чётким при увеличении размера изображения. К сожалению, при работе с цветным фоном результат несколько хуже, чем при работе с белым фоном. Вероятно, это можно исправить путём увеличения объёма набора данных для тренировки, добавлением большего количества цветных изображений и обучением модели в течение большего количества эпох.

Описание модели для снижения степени размытости изображений. Данная модель призвана решить задачу снижения степени размытости изображения. Для этого было принято решение использовать свёрточную нейронную сеть (*Convolutional Neural Network* или *CNN*).

Как и ранее, исходное изображение представлено в формате *RGB*, в связи с чем количество каналов на входе и на выходе нейросети принимается равным трём.

Архитектура данной нейросети схожа с архитектурой ранее рассмотренной сети, но имеет существенные отличия (стрелкой указывается переход от исходного количества каналов к количеству каналов на выходе слоя):

1 *ReflectionPad*: по 4 пикселя снизу, сверху, слева и справа.

2 *Convolution*: $3 \rightarrow 75$, ядро 4×4 , шаг 1.

3 *ReLU*.

4 *Convolution*: $75 \rightarrow 75$, ядро 2×2 , шаг 1.

5 *ReLU*.

6 *Convolution*: $75 \rightarrow 75$, ядро 1×1 , шаг 1.

7 *ReLU*.

8 *Convolution*: $75 \rightarrow 75$, ядро 2×2 , шаг 1.

9 *ReLU*.

10 *Convolution*: $75 \rightarrow 3$, ядро 4×4 , шаг 1.

11 *Sigmoid*.

В архитектуре данной сети используется новый слой: *ReflectionPad*. Необходимость его применения объясняется особенностью слоя свёртки: фактически для каждого канала применяется фильтр, в результате чего значения результирующей матрицы являются суммами произведений элементов фильтра на элементы исходной матрицы. Нетрудно установить, что в случае, когда используется размер фильтра больший, чем 1×1 , результирующая матрица имеет меньший размер, чем исходная.

С целью сохранения исходного размера изображения оно искусственно расширяется на 4 пикселя во все стороны, причём расширяется неслучайно, а путём отражения частей изображения, расположенных на его границе. Для наглядности на рисунке 9 приведён пример использования *ReflectionPad* на 2 пикселя во все стороны для изображения размером 4×4 пикселя.

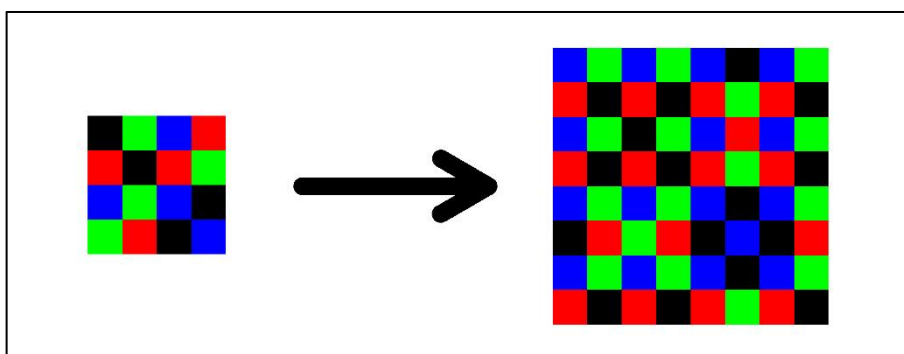


Рисунок 9. Пример использования *ReflectionPad*

Выбор набора данных для обучения модели для снижения степени размытости изображений. На данном этапе было решено использовать некоторую часть из набора данных с *British Machine Vision Conference (BMVC)* [4], а именно сгенерированные функции рассеяния точки.

Функция рассеяния точки (англ. *Point Spread Function* или *PSF*) представляет собой изображение, которое было бы получено при наблюдении точечного источника света в

некоторых условиях. Именно эти условия и можно описать с помощью *PSF*. В частности, такие функции удобно применять при моделировании размытия.

Итак, из набора *BMVC* случайным образом были выбраны функции рассеяния точки. Пример выбранных функций представлен на рисунке 10.

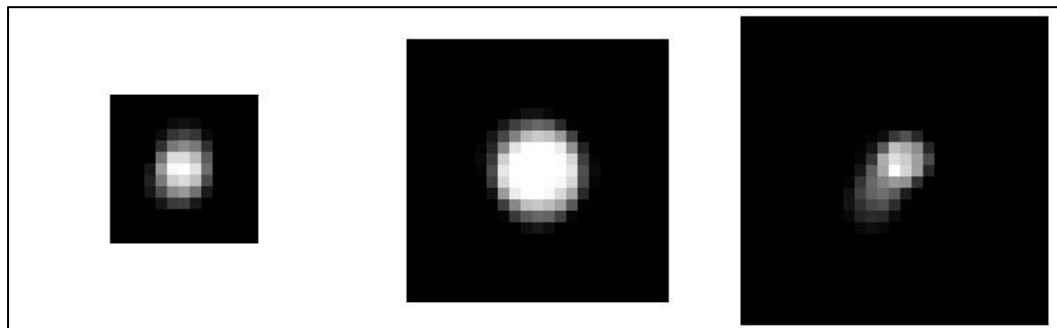


Рисунок 10. Пример используемых *PSF*

Для применения *PSF* достаточно конвертировать рассмотренные представления функции в виде изображения в матрицу, состоящую из нормированных яркостей пикселей, и применить к исходному изображению фильтр, ядром которого выступает указанная матрица.

В качестве исходного был использован набор, аналогичный тому, что использовался для обучения модели для увеличения изображения (см. рисунок 4). К каждому изображению из исходного набора была применена соответствующая ему случайным образом выбранная функция рассеяния точки. Примеры изображений из полученного набора (назовём его «итоговый набор») представлены на рисунке 11.

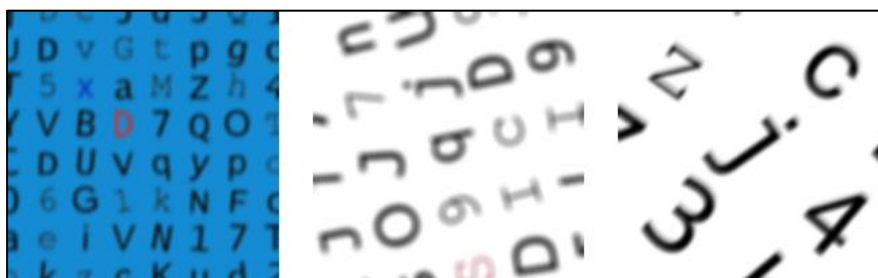


Рисунок 11. Пример изображений итогового набора

Для обучения созданной модели изображения из исходного и итогового наборов были собраны в пары, 70% таких пар отобраны случайным образом непосредственно для обучения модели, остальные 30% используются для валидации результатов. Как и в случае с моделью для увеличения размера изображения, данные не были нормализованы.

Обучение модели для снижения степени размытости изображений и достигнутые результаты. Как и в ранее рассмотренном случае, модель обучалась из начального состояния, которое определялось случайным образом с помощью метода инициализации Ксавье (*Xavier* или *Glorot initialization*) [6].

В качестве меры ошибки была выбрана среднеквадратичная ошибка (*MSELoss*). Для обучения использовался тот же модифицированный алгоритм стохастического градиентного спуска: *Adam*. Аналогично предыдущей модели данный алгоритм показал хорошую скорость сходимости и позволил снизить итоговую ошибку.

Модель обучалась на протяжении 1000 эпох. График зависимости среднеквадратичной ошибки (оранжевым – ошибка на данных для валидации, синим – на

тренировочных данных) от эпохи представлен на рисунке 12.

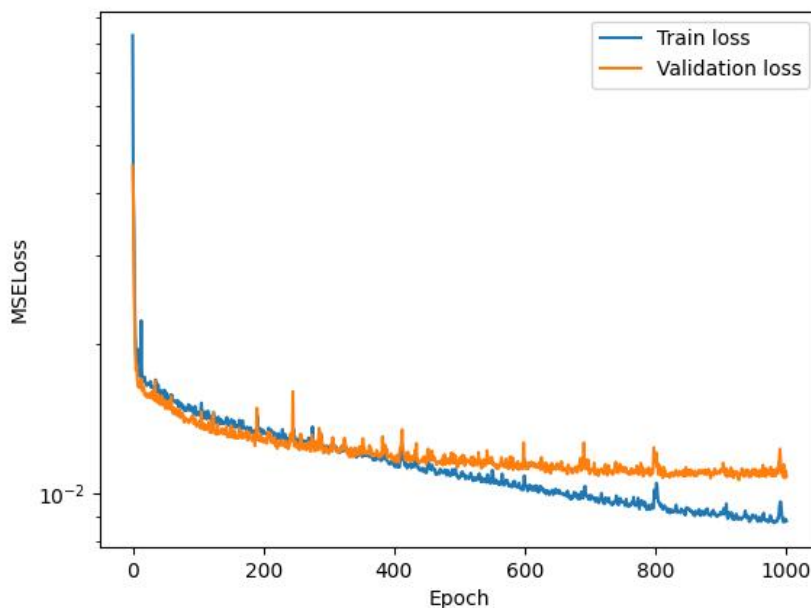


Рисунок 12. Ошибка на валидационных (оранжевым) и тренировочных (синим) данных, логарифмическая шкала

Видно, что ошибка на данных для валидации на последних эпохах практически не уменьшалась, в то время как ошибка на тренировочных данных продолжала падать. Это может свидетельствовать о переобучении модели. Впрочем, нельзя сказать, что ошибка на валидационных данных начала расти, да и, как будет видно далее, нельзя утверждать о плохом в том или ином смысле поведении модели на таких данных.

Пример обработки уменьшенных изображений представлен на рисунке 13, более детальное сравнение – на рисунке 14.

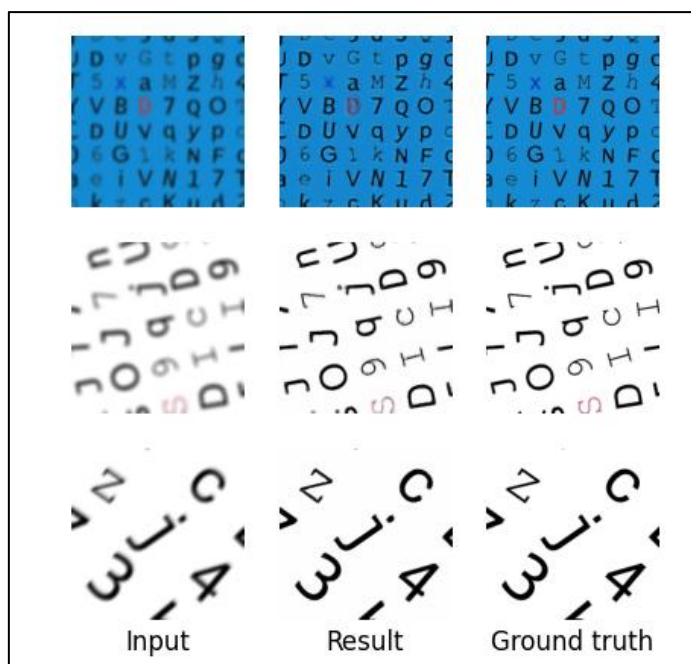


Рисунок 13. Пример обработки размытых изображений

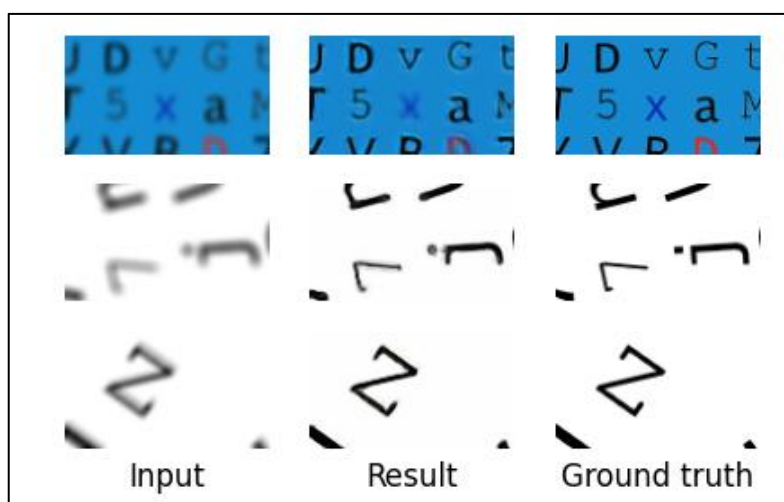


Рисунок 14. Детальное сравнение результатов обработки размытых изображений

Комбинирование обученных моделей. Безусловно, можно применять полученные модели независимо друг от друга, однако зачастую низкое качество изображения подразумевает именно комбинацию размытости и низкого разрешения. По результатам проведённых экспериментов предлагается сначала увеличивать размер изображения, а затем снижать степень его размытости.

Заключение. Основными показателями, влияющими на качество текстовых изображений, являются разрешение и степень размытости. Предложенные алгоритмы успешно решают задачи увеличения размера изображения в 4 раза и снижения степени размытости текста. Примечательно, что рассмотренные нейросети могут обрабатывать изображения практически любых размеров (за исключением самых маленьких). Достигнутые результаты теоретически можно улучшить, например, путём модификации архитектуры нейронных сетей, изменения используемых наборов данных, использования профессионального оборудования для быстрого обучения более глубоких сетей, пред- или постобработки изображений с помощью иных методов.

Список литературы

- [1] X. Xu, D. Sun, J. Pan, Y. Zhang, H. Pfister and M. -H. Yang, "Learning to Super-Resolve Blurry Face and Text Images," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 251-260, doi: 10.1109/ICCV.2017.36.
- [2] Odena, et al., "Deconvolution and Checkerboard Artifacts", Distill, 2016. <http://doi.org/10.23915/distill.00003>.
- [3] Donahue J., Krähenbühl P., Darrell T. Adversarial feature learning //arXiv preprint arXiv:1605.09782. – 2016.
- [4] BMVC OCR test data [Электронный ресурс]. — Режим доступа: https://www.fit.vutbr.cz/~ihradis/CNN-Deblur/BMVC_OCR_test_data.tar.gz. — Дата доступа: 25.11.2023.
- [5] Synthetic Character Set [Электронный ресурс]. — Режим доступа: <https://www.kaggle.com/datasets/shreyasubbu/synthetic-character-set>. — Дата доступа: 10.12.2023.
- [6] Glorot, X. & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, in Proceedings of Machine Learning Research 9:249-256. <https://proceedings.mlr.press/v9/glorot10a.html>.

Авторский вклад

Сачивко Никита Сергеевич – постановка задачи исследования, создание архитектуры нейросетевых моделей, обучение созданных моделей, анализ полученных результатов.

Калугина Марина Алексеевна – руководство исследованием по повышению читабельности текста на изображениях низкого качества с применением нейросетевых моделей.

ALGORITHMS FOR ENHANCING READABILITY OF TEXT IN LOW-QUALITY IMAGES USING DEEP NEURAL NETWORK MODELS

N.S. Sachivko
Student of BSUIR

M.A. Kalugina
Associate Professor of Informatics
Department of the BSUIR

Abstract. This work addresses the problem of unreadable text in low-quality images. Two algorithms have been proposed to improve its readability: an algorithm for upscaling an image and an algorithm for deblurring it. The architecture of neural network models has been described; the results of the proposed algorithms have been presented.

Keywords: machine learning, deep learning, neural networks, neural network models, text recognition, image upscaling, image enhancing, deblurring, enhancing text readability, improving text readability.

УДК 004.023

РАЗРАБОТКА НЕЙРОСЕТЕВОГО МОДУЛЯ ПРОГНОЗИРОВАНИЯ СТРАХОВЫХ ВЫПЛАТ



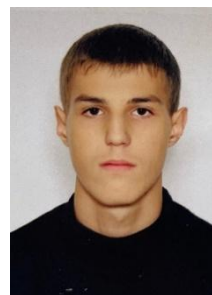
Т.А. Васяева

Декан факультета информационных систем и технологий, доцент кафедры автоматизированных систем управления ДОННТУ, кандидат технических наук, доцент
vasyaeva@gmail.com



Ю.А. Золушкин

Аспирант кафедры автоматизированных систем управления ДОННТУ
illuzium1999@gmail.com



А.В. Яковчук

Магистрант кафедры автоматизированных систем управления ДОННТУ
Yakovchuk-2001@mail.ru

Т.А. Васяева

Окончила Донецкий национальный технический университет. Область научных интересов связана с разработкой методов и алгоритмов искусственного интеллекта, организацией учебного и научно-исследовательского процессов в техническом университете.

Ю.А. Золушкин

Окончил Донецкий национальный технический университет. Область научных интересов связана с разработкой методов машинного обучения, в том числе алгоритмов обработки естественного языка.

А.В. Яковчук

Магистрант Донецкого национального технического университета. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем, машинного обучения.

Аннотация. Решается задача повышения точности планирования финансирования деятельности страхового Фонда, за счет прогнозирования его расходов. Предложено использовать нейросетевой подход для прогнозирования расходов Фонда. Выполнен предварительный анализ имеющихся статистических данных. Предложена нейросетевая модель для прогнозирования страховых выплат. Проведены эксперименты по подбору гиперпараметров нейронной сети.

Ключевые слова: нейронная сеть, временной ряд, автокорреляционная функция, социальное страхование.

Введение. Социальное страхование сегодня, в современном обществе, играет большую роль в функционировании государства и экономики, а также в поддержке жизненного уровня населения. Социальное страхование является одним из институтов экономического развития общества, независимо от его политического устройства и статуса на мировой арене. Основа развития страхового рынка заключается в необходимости обеспечения бесперебойности воспроизведенного процесса путем

оказания денежной помощи пострадавшим в случае непредвиденных неблагоприятных обстоятельств.

Обязательное условие страхового рынка - наличие общественной потребности на страховые услуги и наличие страховщиков, способных удовлетворить эти потребности.

В период не стабильного развития экономики часто возникают ситуации, когда происходит неправильное и не точное планирования бюджетных средств, что влечет за собой нехватку финансирования для выполнения обязательств. Для того, чтобы предугадать поведение внутри финансирования страхового Фонда осуществляют прогнозирование финансовых показателей [1].

Анализ предметной области и цель исследования. Фонд социального страхования от несчастных случаев на производстве и профессиональных заболеваний Донецкой Народной Республики (далее – Фонд) является некоммерческой самоуправляющейся организацией, которая осуществляет профилактическую работу по обеспечению безопасных условий труда на предприятиях и учреждениях республики и управляет средствами государственного социального страхования от несчастных случаев на производстве и профессиональных заболеваний.

Для исследования было выбрано отделение Фонда социального страхования от несчастных случаев на производстве и профессиональных заболеваний Донецкой Народной Республики в Кировском районе г. Макеевки.

Отделение Фонда обязано произвести прогнозирование и планирование основных расходов на месяц, квартал, год и передать данные в Центральный аппарат. Центральный аппарат Фонда в свою очередь, выделяет бюджетные средства для функционирования деятельности отделения, которое в дальнейшем производит выплаты и закупку средств медицинской реабилитации потерпевшим, которые состоят на учете в Фонде.

Основными видами расходов отделения Фонда являются:

- страховая выплата в случае частичной или полной утраты трудоспособности;
- расходы на лекарственные средства и изделия медицинского назначения;
- расходы на поддержание существования Фонда (аренда, коммунальные услуги, зарплата и т.д.).

От правильности и точности прогнозирования и планирования зависит объем финансирования, которое выделит Центральный аппарат, поэтому погрешность должна быть минимальной.

Целью исследования является повышение точности планирования финансирования деятельности страхового Фонда, за счет прогнозирования его расходов.

Постановка задачи и данные для исследования. В качестве исследуемых данных в условиях отделения Фонда были собраны статистические данные за 3 года, которые представляют собой суммы ежемесячных страховых выплат; ежемесячных выплат по медицине; ежемесячных расходов на содержание Фонда в российских рублях. Собранные параметры представляют собой одномерные временные ряды с интервалом в один месяц.

Наша задача сводится к прогнозированию одномерных временных рядов. Задача прогнозирования временных рядов заключается в следующем. Заданы дискретные отсчеты $\{y(t_1), y(t_2), \dots, y(t_n)\}$ в последовательные моменты времени t_1, t_2, \dots, t_n . Необходимо рассчитывать значения $y(t_{n+1})$ в некоторый будущий момент времени t_{n+1} .

Анализ данных. При анализе и изучении временного ряда аналитик должен на основе некоторого отрезка ряда конечной длины сделать выводы о характере и закономерностях процесса, который описывается данным рядом. Таким образом, цель анализа временного ряда – описание характерных особенностей ряда, для дальнейшего построения модели ряда для предсказания будущих значений на основе прошлых наблюдений [2].

Для этого будем использовать аппарат корреляционного анализа. Корреляция – взаимосвязь двух или более величин, при этом изменения значений одной из величин сопровождается изменению значений другой величины. При анализе временных рядов вычисляется автокорреляция ряда и строится автокорреляционная функция [3].

Таблица 1. Данные для расчета автокорреляционной функции (АКФ) на примере данных по ежемесячным выплатам по медицине

X	355775	361541	352862	342719	...	339457	363758	376256
X_{i-1}		355775	361541	352862	...	350852	339457	363758
X_{i-2}			355775	361541	...	331852	350852	339457
X_{i-3}				355775	...	328785	331852	350852
X_{i-4}					...	356581	328785	331852
...
X_{i-n}					...			355775

Коэффициент автокорреляции изменяется в диапазоне $[-1;1]$, где $r_k = 1$ означает полную корреляцию. Рассчитав коэффициенты автокорреляции для каждого сдвига, получим автокорреляционную функцию (АКФ). Автокорреляционная функция – последовательность коэффициентов автокорреляции уровней первого, второго и т. д. порядков. Расчет АКФ на примере данных по ежемесячным выплатам по медицине, для ежемесячных страховых выплат и ежемесячных расходов на содержание Фонда в табл. 2-4 и графики АКФ на рис. 1.

Таблица 2. Результаты АКФ по ежемесячным страховым выплатам.

k	0	1	2	3	4	5	6	7	8	9	10	11
r_k	1	0,99	0,98	0,98	0,98	0,97	0,97	0,97	0,96	0,97	0,973	0,98
k	12	13	14	15	16	17	18	19	20	21	22	23
r_k	0,98	0,98	0,98	0,98	0,97	0,95	0,93	0,921	0,914	0,91	0,90	0,90
k	24	25	26	27	28	29	30	31	32	33	34	35
r_k	0,92	0,92	0,93	0,94	0,94	0,92	0,84	0,761	0,72	0,67	0,62	0,67

Таблица 3. Результаты АКФ по выплатам по медицине

k	0	1	2	3	4	5	6	7	8	9	10	11
r_k	1	0,07	0,07	0,5	-0,3	-0,5	0,3	-0,6	-0,3	0,4	-0,1	-0,00
k	12	13	14	15	16	17	18	19	20	21	22	23
r_k	0,92	0,07	0,12	0,52	-0,26	-0,5	0,26	-0,64	-0,34	0,35	-0,15	-0,05
k	24	25	26	27	28	29	30	31	32	33	34	35
r_k	0,89	0,02	0,13	0,6	-0,27	-0,45	0,28	-0,7	-0,46	0,36	-0,27	-0,13

Таблица 4. Результаты АКФ по расходам на содержание Фонда

k	0	1	2	3	4	5	6	7	8	9	10	11
r_k	1	-0,04	0,15	0,03	-0,19	-0,34	0,02	-0,41	-0,28	0,098	0,056	0,07
k	12	13	14	15	16	17	18	19	20	21	22	23
r_k	0,78	-0,1	0,08	0,16	-0,13	-0,15	-0,05	-0,48	-0,22	-0,02	0,04	0,14
k	24	25	26	27	28	29	30	31	32	33	34	35
r_k	0,72	-0,09	0,35	0,19	-0,09	-0,18	-0,24	-0,54	-0,19	-0,12	0,06	0,07

Проанализировав полученные результаты, можно сделать вывод, что временной ряд по ежемесячным страховым выплатам содержат тренд, так как степень взаимной зависимости между последовательными значениями ряда и корреляция между ними очень высоки. При этом коэффициент автокорреляции значителен для многих сдвигов ряда.

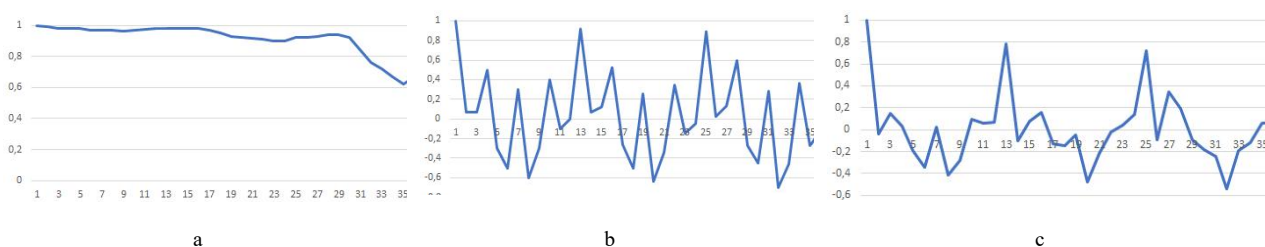


Рисунок 1. Автокорреляционная функция временного ряда *a*) по ежемесячным страховым выплатам, *b*) по выплатам по медицине, по расходам на содержание Фонда

Скорее всего, временной ряд по выплатам по медицине и по расходам на содержание Фонда содержат сезонную компоненту, так как коэффициент автокорреляции периодически увеличивается. Наличие сезонной компоненты можно предположить и по визуальному анализу графиков временных рядов, но так как данные всего за три года, а предположительный период равен году, то недостаточно данных для однозначного утверждения наличия сезонной компоненты.

Учитывая наличие тренда и высоких значений коэффициентов автокорреляции для многих сдвигов ряда, можно предположить, что для построения прогностической модели можно брать порядка 10 предшествующих значений для временного ряда по ежемесячным страховым выплатам. Но учитывая ограниченное количество данных, остановимся на 5 входах, так как увеличение числа входов существенно сократит обучающую выборку. По выплатам по медицине и по расходам на содержание Фонда будем так же использовать 5 входов.

Обобщенная модель и построение обучающей выборки. Обобщенная модель прогнозирования временного ряда, имеет следующий вид (рис. 2).

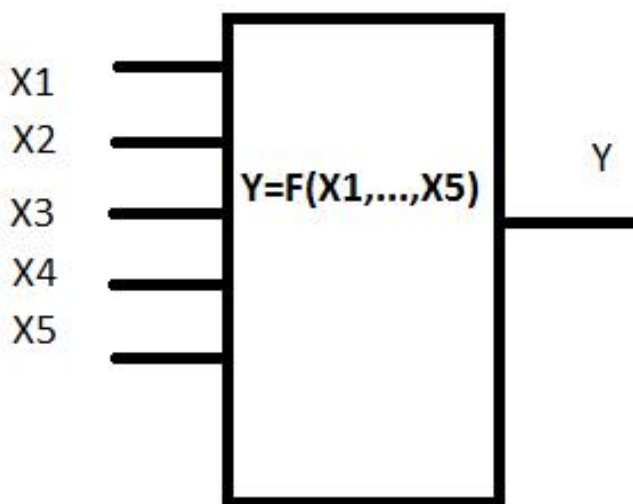


Рисунок 2. Обобщенная модель для прогнозирования временного ряда по ежемесячным страховым выплатам

Обучающая выборка строится методом скользящего окна [3], который широко применяется при работе с временными рядами. Обучающие выборки для такой модели будут иметь такой формат (табл. 5).

Разработка нейросетевой модели и ее оптимизация. Для прогнозирования временных рядов, как показывает опыт, хорошо показывают себя многослойные

нейронные сети. Можно также предположить, что 3-х слоев будет достаточно, затем проверить и подтвердить экспериментально. На выходном слое будет один нейрон, так как прогнозируем на один шаг (месяц) вперед. Количество нейронов на скрытых слоях будем определять экспериментально. К другим гиперпараметрам модели относят: количество эпох обучения, функции активации нейронов, функция ошибок. Так же будем экспериментально подбирать оптимизатор.

Таблица 5. Фрагмент обучающей выборки по ежемесячным страховым выплатам

X_1	X_2	X_3	X_4	X_5	Y
1901542	1913564	1913564	1926764	1982457	1984254
1913564	1913564	1926764	1982457	1984254	1995424
1913564	1926764	1982457	1984254	1995424	1997403
...
2265412	2267521	2300452	2357521	2378552	2378552

В среде *Colab Laboratory* выполнена программная реализация нейронной сети для прогнозирования. Для моделирования использовался язык *Python*, с применением библиотек *pandas*; *numpy*; *keras* (в качестве основы *tensorflow*).

Colab Laboratory [5] – это бесплатная среда для ноутбуков *Jupyter*, которая не требует настройки и полностью работает в облаке. *Colab Laboratory* позволяет писать и выполнять код, а также получать доступ к мощным вычислительным ресурсам, что является важным преимуществом при работе с ресурсоемкими алгоритмами машинного обучения.

Keras [6] – одна из самых мощных и простых в использовании библиотек *Python* для разработки и оценки моделей глубокого обучения, охватывает эффективные библиотеки численных вычислений *Theano* и *TensorFlow*. Преимущество этого состоит главным образом в том, есть возможность работать с нейронными сетями достаточно простым способом.

Экспериментальные исследования. Проведем эксперименты по выбору архитектуры НС и настройки основных гиперпараметров модели.

Для начала проведем исследования для данных ежемесячных страховых выплат. Структура сети будет следующая: 5 входов (согласно данным), 5 нейронов на первом слое, 6 нейронов на втором слое и 1 нейрон на выходном слое, функции активации: *relu*, *relu*, *linear*. В таблице 6 представлены эксперименты на обучающей, валидационной, тестовой выборках при разном количестве эпох. Точность анализируем по *Mean Squared Error (MSE)* и *Mean absolute error (MAE)*. *MSE* рассчитывается как среднее значение квадратов разностей между прогнозируемыми и фактически наблюдаемыми значениями. *MAE* рассчитывается как среднее значение абсолютной разницы между прогнозами и фактическими наблюдениями. Оптимизатор будем использовать *adam*. Результаты экспериментов представлены в таблице 6. Как видно из таблицы 6, на 500 эпохе наблюдаем переобучение сети, следовательно дальнейшее обучение нецелесообразно.

Далее, используя ту же архитектуру, проведем эксперименты по ежемесячным выплатам по медицине. Результаты экспериментов представлены в таблице 7. Анализируя таблицу 7, можно заметить, что при модели обучается плохо. Уже на 100 эпохе появляются признаки переобучения, далее модель обучается незначительно, а при обучении на 500 эпохе опять наблюдаем признаки переобучения.

Таким же образом, проведем исследования на данных по ежемесячным расходам на содержание Фонда. Результаты экспериментов представлены в таблице 8. Сеть практически не обучается.

Изменим архитектуру нейронной сети на следующую: 5 входов, 5 нейронов на

первом слое, 6 нейрона на втором слое, 1 на выходном; функции активации: гиперболический тангенс, sigmoid, relu. Результаты экспериментов представлены в таблицах 9-11.

Таблица 6 Ошибка обучения на НС по ежемесячным страховым выплатам

epochs	Training set		Validation set		Test set	
	mse	mae	mse	mae	mse	mae
25	0.0194	0.1099	0.2645	0.5118	0.0053	0.6201
50	0.0167	0.1029	0.2253	0.4713	0.5226	0.7126
75	0.0032	0.0488	0.0292	0.1675	0.4930	0.6937
100	0.0012	0.0290	0.0050	0.0628	0.0607	0.2383
200	0.0016	0.0307	0.0049	0.0543	0.0060	0.0658
300	0.0007	0.0192	0.0014	0.0292	0.0060	0.0603
400	0.0013	0.0280	0.0010	0.0296	0.0014	0.0359
500	0.0008	0.0192	0.0033	0.0511	0.0042	0.0557

Таблица 7 Ошибка обучения на НС по ежемесячным выплатам по медицине

epochs	Training set		Validation set		Test set	
	mse	mae	mse	mae	mse	mae
25	0.1972	0.3793	0.2299	0.4109	0.2058	0.3852
50	0.1331	0.3138	0.1889	0.3851	0.2979	0.4668
75	0.0232	0.1262	0.0267	0.0906	0.2057	0.3875
100	0.0577	0.2014	0.0754	0.2437	0.0509	0.1988
200	0.0235	0.1226	0.0277	0.1164	0.0917	0.2566
300	0.0200	0.1161	0.0241	0.1207	0.0511	0.2190
400	0.0187	0.1192	0.0331	0.1341	0.0454	0.1878
500	0.0544	0.1887	0.0838	0.2270	0.0429	0.1920

Таблица 8 Ошибка обучения на НС по ежемесячным расходам на содержание Фонда

epochs	Training set		Validation set		Test set	
	mse	mae	mse	mae	mse	mae
25	0.1038	0.2661	0.2627	0.4759	0.0426	0.1632
50	0.1273	0.2808	0.1323	0.3163	0.0358	0.1597
75	0.0551	0.1954	0.1352	0.3215	0.0620	0.2016
100	0.0607	0.2085	0.0904	0.2693	0.0404	0.1739
200	0.0592	0.1913	0.0807	0.2215	0.0429	0.1668
300	0.0507	0.1945	0.0946	0.2600	0.0478	0.1871
400	0.0483	0.1855	0.0982	0.2451	0.0536	0.1901
500	0.0223	0.1014	0.1322	0.3145	0.0543	0.1845

Как видно из таблицы 9, на 500 эпохе наблюдаем переобучение сети, следовательно дальнейшее обучение нецелесообразно. Так же отмечаем, что нелинейные функции активации на скрытых слоях дают лучшие результаты. Изменим оптимизатор на *adamax* и проведем эксперименты (табл. 10.), а затем применим оптимизатор *SGD* (табл. 11).

Анализируя данные, можно сказать, что оптимизатор *SGD* показывает лучшие результаты. При этом следует отметить, что сеть имеет склонность к переобучению.

Аналогичным образом был выполнен еще ряд экспериментов. Отмечено, что увеличение количества нейронов и / или количество слоев приводит к переобучению при меньшем количестве эпох, но не способствуют улучшению результата (снижению ошибки).

Таблица 9 Ошибка обучения на НС оптимизатор *adam*

<i>epochs</i>	<i>Training set</i>		<i>Validation set</i>		<i>Test set</i>	
	<i>mse</i>	<i>mae</i>	<i>mse</i>	<i>mae</i>	<i>mse</i>	<i>mae</i>
25	0.0420	0.0236	0.0203	0.0134	0.0094	0.3545
50	0.0210	0.0228	0.0148	0.0148	0.1430	0.1213
75	0.0530	0.0537	0.03523	0.1571	0.1967	0.1039
100	0.0160	0.0259	0.03523	0.0852	0.1967	0.2648
200	0.0148	0.0176	0.1375	0.0645	0.1967	0.2849
300	0.0170	0.0196	0.0714	0.0215	0.0194	0.2460
400	0.0162	0.0147	0.0950	0.0148	0.1967	0.1319
500	0.0168	0.0139	0.0600	0.0148	0.1967	0.1039

Таблица 10 Ошибка обучения на НС оптимизатор *adamax*

<i>epochs</i>	<i>Training set</i>		<i>Validation set</i>		<i>Test set</i>	
	<i>mse</i>	<i>mae</i>	<i>mse</i>	<i>mae</i>	<i>mse</i>	<i>mae</i>
25	6.9462e-05	0.0135	0.1039	0.2313	0.1967	0.4219
50	0.0002	0.0135	0.0497	0.0572	0.0045	0.1177
75	6.7806e-05	0.0238	0.1039	0.2313	0.1967	0.1180
100	3.5834e-05	0.0120	0.1039	0.2313	0.1967	0.1185
200	0.0006	0.0057	0.1039	0.2313	0.1967	0.0755
300	1.9409e-05	0.0232	0.1039	0.2313	0.1967	0.3214
400	6.1079e-05	0.0032	0.1039	0.2313	0.1967	0.4219
500	0.0002	0.0025	0.0331	0.0062	0.0012	0.4219

Таблица 11 Ошибка обучения на НС оптимизатор *SGD*

<i>epochs</i>	<i>Training set</i>		<i>Validation set</i>		<i>Test set</i>	
	<i>mse</i>	<i>mae</i>	<i>mse</i>	<i>mae</i>	<i>mse</i>	<i>mae</i>
25	0.0006	0.0020	0.1101	0.0148	0.1751	0.0208
50	0.0032	0.0019	0.0538	0.0148	0.1297	0.0210
75	0.0004	0.0023	0.0538	0.0032	0.0261	0.0211
100	0.0004	0.0023	0.0538	0.0148	0.0148	0.1967
200	0.0005	0.0017	0.0038	0.0148	0.1571	0.0205
300	0.0004	0.0016	0.0538	0.0148	0.0852	0.0205
400	0.0005	0.0015	0.0538	0.0148	0.0645	0.0210
500	0.0005	0.0022	0.0538	0.0148	0.0215	0.0094

Заключение. В данной работе предложен подход к прогнозированию основных видов расходов отделения Фонда. Предложено использовать нейросетевую модель для прогнозирования временных рядов страховых выплат. Проведены эксперименты по разработке нейронной сети согласно которым можно сделать вывод: рекомендовано использовать нелинейные функции активации на скрытых слоях и *relu* на выходном слое; оптимизатор *SGD*; достаточно использовать 3 слоя нейронной сети. Разработана архитектура НС: 5 входов, 5 нейронов на первом слое, 6 нейрона на втором слое, 1 на выходном; функции активации: гиперболический тангенс, *sigmoid*, *relu*, оптимизатор *SGD*.

Список литературы

[1]. Васяева Т.А., Шуватова Е.А. Яковчук А.В. Анализ методов прогнозирования для управления финансовыми средствами страховой компании. Информатика, управляющие системы, математическое и компьютерное моделирование» (ИУСМКМ-2023): сборник трудов XIV международной научно-технической

конференции в рамках IX Международного Научного форума Донецкой Народной Республики / Под ред. Р.В. Мальчевой, И.В. Матях; ФГБОУ ВО «ДонНТУ». – Донецк: ДонНТУ, 2023.

[2]. Peter J. Brockwell, Richard A. Davis: Introduction to Time Series and Forecasting. Second Edition. Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA (2002).

[3]. Sergii, K., Yurii, S., Tatyana, V., Natalia, A.: Feature Selection for Time-Series Prediction in Case of Undetermined Estimation. In: Samsonovich A., Klimov V., Rybina G. (eds.) Biologically Inspired Cognitive Architectures (BICA) for Young Scientists, Advances in Intelligent Systems and Computing, vol. 449, pp. 85-97. Springer, Cham (2016). DOI: 10.1007/978-3-319-32554-5_12

[4]. Vasyaeva, T., Martynenko, T., Khmilovyi, S., Andrievskaya, N. (2019). Stock Prices Forecasting with LSTM Networks. In: Kuznetsov, S., Panov, A. (eds) Artificial Intelligence. RCAI 2019. Communications in Computer and Information Science, vol 1093. Springer, Cham. https://doi.org/10.1007/978-3-030-30763-9_5

[5]. Google Colaboratory. - <https://colab.research.google.com/notebooks/welcome.ipynb>.

[6]. Keras: The Python Deep Learning library. - <https://keras.io>.

Авторский вклад

Васяева Татьяна Александровна – руководство исследованием по прогнозированию основных видов расходов отделения Фонда социального страхования.

Золушкин Юрий Алексеевич – реализация алгоритма и разработка нейросетевой модели прогнозирования временных рядов.

Яковчук Андрей Владимирович – тестирование нейросетевого модуля прогнозирования страховых выплат.

DEVELOPMENT OF A NEURAL NETWORK MODULE FOR PREDICTING INSURANCE PAYMENTS

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

Yu.A. Zolushkin

Postgraduate student of the department automated control systems of DONNTU

A.V. Yakovchuk

Master's student of the department automated control systems of DONNTU

Abstract. The problem of improving the accuracy of planning the funding of the activities of the Insurance Fund by forecasting its expenses is being addressed. A neural network approach is proposed for predicting the Fund's expenses. A preliminary analysis of the existing statistical data has been carried out. A neural network model for predicting insurance payouts has been proposed. Experiments have been conducted to tune the hyperparameters of the neural network.

Keywords: neural network, time series, autocorrelation function, social insurance.

УДК 519.67

ПРОГНОЗИРОВАНИЕ ВРЕМЕННОГО РЯДА В PYTHON С ПОМОЩЬЮ МЕТОДА ХОЛЬТА-УИНТЕРСА



С.К. Корсакова

Студентка группы 121701,
БГУИР
swe123ta@gmail.com



Н.В. Русина

Старший преподаватель
кафедры экономической
информатики БГУИР
rusina@bsuir.by



В.М. Бондарик

Декан факультета
доуниверситетской
подготовки и
профессиональной
ориентации БГУИР,
кандидат технических наук,
доцент
bondarik@bsuir.by

С.К. Корсакова

Является студенткой группы 121701 факультета информационных технологий и управления Белорусского государственного университета информатики и радиоэлектроники.

Н.В. Русина

Окончила Могилёвский государственный университет им. А. Кулешова. Область научных интересов связана с разработкой информационной системы электронного зачисления в ВУЗы, организацией учебного и научно-исследовательского процессов в техническом университете.

В.М. Бондарик

Окончил Минский радиотехнический институт. Область научных интересов связана с разработкой медицинской электронной техники, организацией учебного и научно-исследовательского процессов в техническом университете.

Аннотация. Прогнозирование – один из способов планирования, который позволяет провести подготовку к ситуациям, которые могут возникнуть в будущем, помогая тем самым оптимизировать расходы ресурсов: временных, материальных и т.д. Прогнозирование может применяться в любой социальной сфере: в экономике, образовании, медицине, бизнесе, на производстве и т.д. В данной статье рассматривается один из способов прогнозирования с помощью временного ряда методом Хольта-Уинтерса.

Ключевые слова: временной ряд, метод Хольта-Уинтерса, прогнозирование.

Введение. Прогнозирование является важнейшим аспектом процесса стратегического управления, поскольку позволяет ставить реалистичные цели, принимать эффективные решения и оптимизировать работу организации с помощью стратегически полезных инструментов. Социально-экономические преобразования можно осуществить с использованием методов методического прогнозирования в стратегии управления, что значимо в рыночных условиях.

Цель прогноза – получение научно-обоснованных вариантов тенденций развития показателей качества и затрат, используемых при разработке перспективных планов и

проведении научно-исследовательских и опытно-конструкторских работ и всей системы менеджмента.

Прогнозирование предполагает обоснованные предположения о будущих событиях, которые могут повлиять на развитие предприятия. Предприятия могут прогнозировать продажи, финансы, потребительский спрос и рыночные изменения, изучая прошлые данные, тенденции и закономерности. Прогнозирование помогает принимать решения, планировать и управлять рисками. Методы прогнозирования обычно включают анализ больших массивов данных и использование статистических моделей. Прогнозы обычно носят вероятностный характер. Непредвиденные события могут повлиять на прогнозы, поэтому важно регулярно просматривать и обновлять их по мере поступления новой информации.

При построении прогноза обычно используют математические методы для анализа больших массивов исходных данных, в качестве которых часто выступают временные ряды.

Временной ряд – это последовательность значений, которые определяются в определённом временном промежутке. Временные ряды удобно использовать для анализа и прогноза в том случае, когда необходимо определить, что будет происходить в определённый временной интервал (час, день, месяц или год) [1].

Прогнозирование временных рядов – это процесс анализа данных временных рядов с использованием статистики и моделирования для последующего прогнозирования и принятия важных решений.

Временные ряды делятся на детерминированные и недетерминированные. В первом случае временной ряд может быть выражен математической формулой, что упрощает анализ и прогнозирование. Для недетерминированных временных рядов прогнозирование усложняется, так как анализ и прогнозирование происходит по средним значениям и дисперсии [1].

Временные ряды могут быть стационарными или нестационарными. В стационарных временных рядах статистические свойства не зависят от времени и, следовательно, их результат легко спрогнозировать. В то время как в нестационарных рядах статистические свойства меняются со временем, поэтому и результат предсказать сложно [1].

Использование различных математических методов для анализа временных рядов позволяет автоматизировать процесс прогнозирования.

Описание метода Хольта-Уинтерса и реализация его на языке программирования Python. В статье представлена реализация алгоритма прогнозирования с использованием метода Хольта-Уинтерса на языке программирования Python. В качестве организации, для которой строится прогноз развития, выбрано физиотерапевтическое отделение поликлиники. Проводится анализ эффективности функционирования физиотерапевтического отделения поликлиники с учетом временных и технических возможностей физиотерапевтического оборудования, его пропускной способности, оптимального использования человеческого ресурса медицинского персонала.

Метод Хольта-Уинтерса (тройное сглаживание) принимает в расчет изменение тренда и колебания значений остатков вокруг этого тренда, а также дополнительно учитывает сезонные колебания и является развитием метода Хольта (двойное сглаживание) [2, 3].

Временной ряд можно разложить на 3 составляющие: тренд, сезонность и случайная составляющая. Модель Хольта-Уинтерса использует идеи экспоненциального сглаживания, но при этом является более сложной и может применяться к рядам, содержащим тенденцию и сезонность. Данная модель работает по достаточно сложным

математическим формулам. Эту модель предложено реализовать в одной из библиотек *Python*, где не требуется программировать все формулы самостоятельно.

Прогнозирование реализовано с использованием метода *statsmodels.tsa.holtwinters.ExponentialSmoothing()*. В данном методе, в зависимости от наличия во временном ряду тренда и сезонности, добавляются или убираются параметры *trend* и *seasonal*, а также определяют периодичность сезонности параметром *seasonal_periods*.

Для реализации модели были использованы следующие библиотеки: *pandas* для упрощения работы с табличными данными, *numpy* для математических операций, *statsmodels* для разбиения временного ряда на компоненты и последующего прогнозирования, а также *sklearn* для оценки построенной модели.

Выполнено прогнозирование методом Хольта-Уинтерса путем обработки массива данных о ежемесячном количестве пациентов физиотерапевтического отделения поликлиники в период с 2009 по 2020 годы (таблица 1).

Таблица 1. Часть исходного массива данных (всего 144 строки)

№	Month	Patients	Month*	#Patients
0	2009-01	112	2009-01-01	4.718499
1	2009-02	118	2009-02-01	4.770685
2	2009-03	132	2009-03-01	4.882802
3	2009-04	129	2009-04-01	4.859812
4	2009-05	121	2009-05-01	4.795791
5	2009-06	135	2009-06-01	4.905275
6	2009-07	148	2009-07-01	4.997212
7	2009-08	148	2009-08-01	4.997212
8	2009-09	136	2009-09-01	4.912655
9	2009-10	119	2009-10-01	4.779123
10	2009-11	104	2009-11-01	4.644391
...				

При реализации алгоритма на языке программирования *Python* заменили индексы данных на *Datetime* (см. таблицу 1, столбец *Month**):

```
data['Month'] = pd.to_datetime(data['Month'])
data = data.set_index('Month')
```

Для исключения вероятности применения неправильной модели и с целью применения аддитивной модели разложения временного ряда прологарифмировали данные столбца *Patients* (см. таблицу 1, столбец *#Patients*).

По результатам моделирования с использование разработанного на языке программирования *Python* программного обеспечения и временного ряда (см. таблицу 1) получили графики составляющих временного ряда (тренд, сезонность и случайные колебания) (рисунок 1).

По полученным компонентам можно определить вид ряда с точки зрения стационарности. В стационарном ряду такие компоненты как тренд и сезонность отсутствуют, а в нестационарном – есть. Определение наличия или отсутствия этих компонент необходимо для выбора вида модели и последующего моделирования. Анализ графиков на рисунке 1 позволяет сделать вывод о том, что используемый временной ряд нестационарный.

Для построения модели Хольта-Уинтерса и прогнозирования временного ряда разделили имеющийся массив данных из 144 наблюдений на обучающую и тестовую части по 124 и 20 строки соответственно.

Последующие тренировка и прогнозирование позволили визуализировать результат моделирования (рисунок 2) и доказали, что полученное предсказание соответствует действительности.

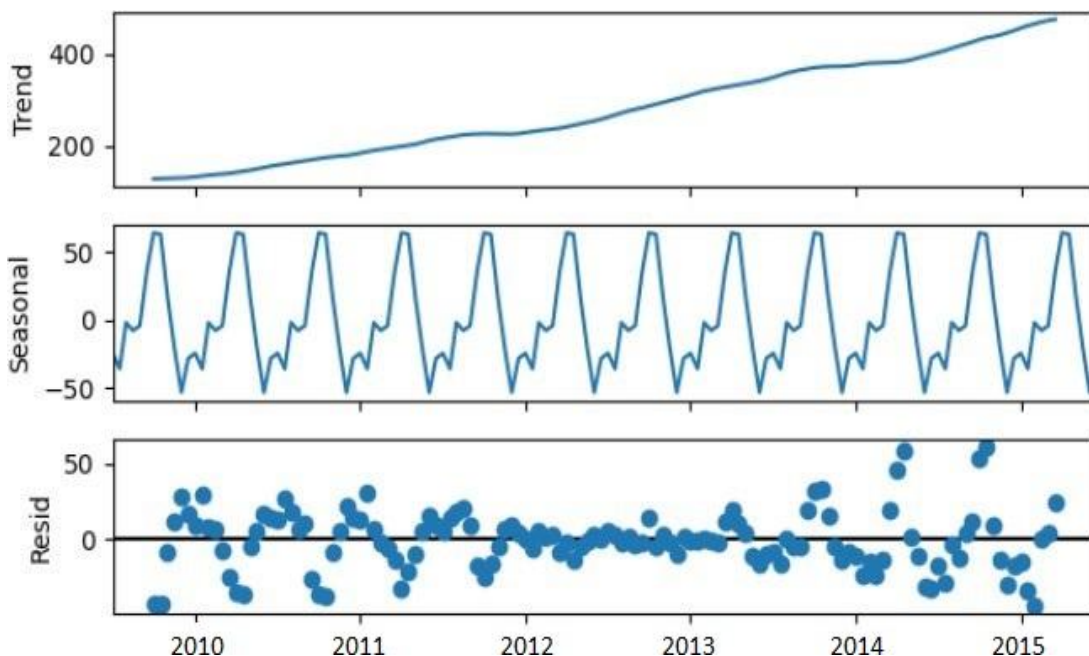


Рисунок 1. Тренд, сезонность и случайные колебания

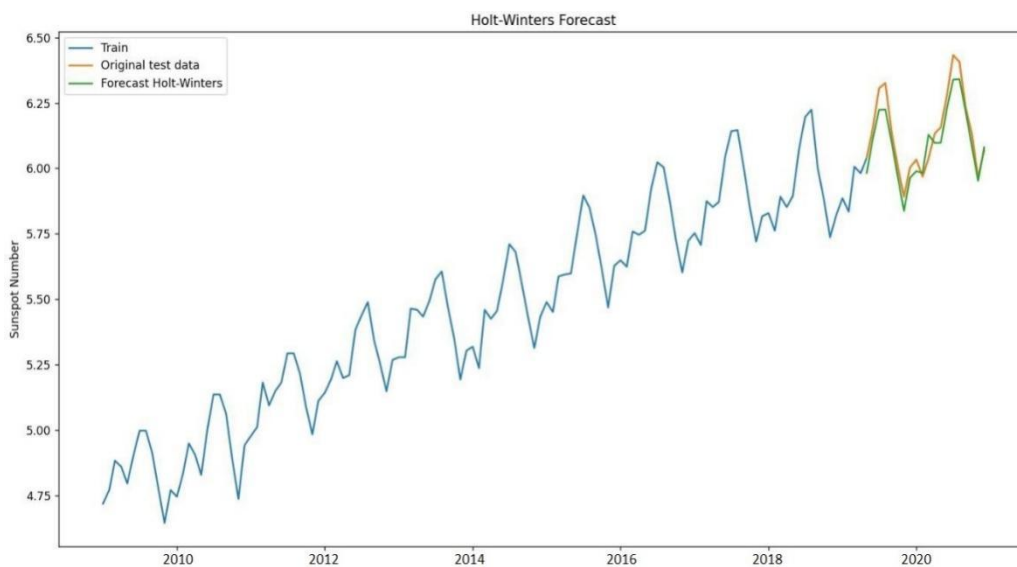


Рисунок 2. Визуализация процесса обучения и прогнозирования

Далее провели оценку результатов прогнозирования, которая включает в себя оценку остатков, определение среднеквадратичной ошибки и коэффициента детерминации (рисунок 3).

Анализ результатов моделирования показал, что систематические отклонения остатков незначительно отклонялись от нуля, что свидетельствует о том, что модель хорошо прогнозирует данные.

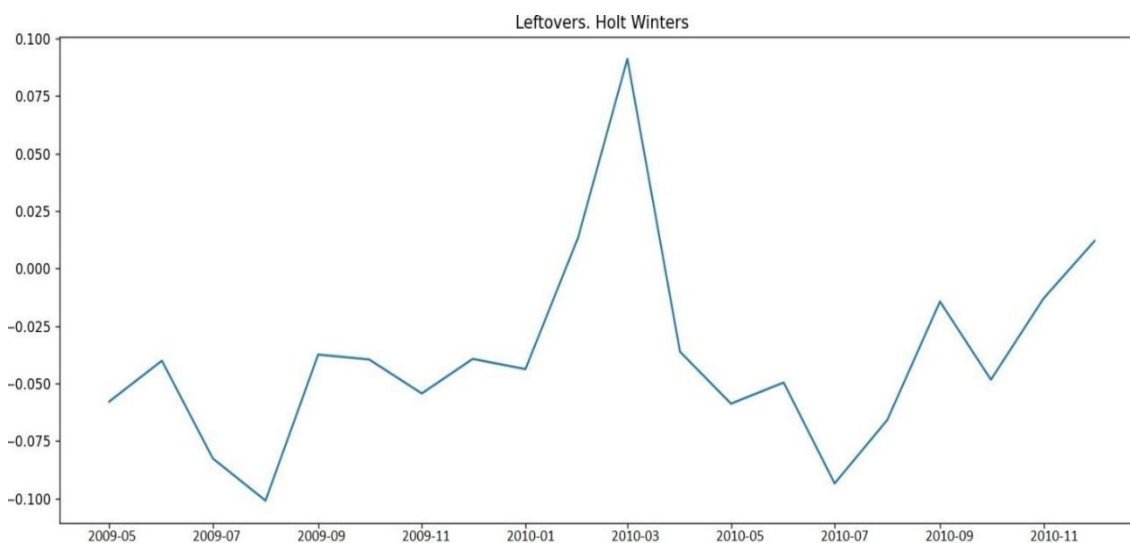


Рисунок 3. Оценка результатов прогнозирования

Рассчитанная по предложенной модели среднеквадратичная ошибка (*Square error of Holt Winter predictions*) оказалась равна 0.0031541506272186695. Полученная ошибка говорит о том, что спрогнозированные значения не сильно отличаются от эталонных значений. Рассчитанный по предложенной модели коэффициент детерминации (*Determination coefficient for Holt Winters*) оказался равен 0.8572937678697624. Полученный коэффициент показывает, что дисперсия модели достаточно мала и модель хорошо описывает данные.

Таким образом, предложенная модель адекватна и имеет высокую значимость.

Заключение. Предложенный для прогнозирования процессов с длительными временными интервалами метод Хольта-Уинтерса учитывает одновременно несколько факторов и строит прогнозы на большой промежуток времени, учитывая значительные неравномерности в распределении значений временных рядов. Применение предложенной модели позволяет заранее спланировать график работы медицинского персонала физиотерапевтического отделения поликлиники, оптимизировать затраты на использование и обслуживание оборудования, не теряя при этом качество обслуживания.

Прогнозирование с помощью метода Хольта-Уинтерса позволяет правильно устанавливать реалистичные цели и задачи для своей деятельности. Прогнозируя результаты развития организации, позволяют устанавливать достижимые цели и отслеживать их прогресс с течением времени.

Список литературы

[1] Лукашин, Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов. Москва : Финансы и статистика, 2003. – 416 с.

[2] Светуных, И. С. Методы социально-экономического прогнозирования в 2 т. Т. 2 модели и методы : учебник и практикум для вузов / И. С. Светуных, С. Г. Светуных. — Москва : Издательство Юрайт, 2024. — 447 с.

[3] Эконометрика. Учебник / Под ред. Елисейевой И. И. – 2-е изд. – Москва : Финансы и статистика, 2006. – 576 с.

Авторский вклад

Корсакова Светлана Константиновна – программирование в *Python* метода Хольта-Уинтерса, тестирование системы, получение и обработка результатов исследования.

Русина Наталья Владимировна – постановка задачи исследования, описание принципа работы метода Хольта-Уинтерса, анализ полученных результатов.

Бондарик Василий Михайлович – руководство исследованием по решению задачи прогнозирования с помощью метода Хольта-Уинтерса, формирование структуры статьи.

TIME SERIES PREDICTION IN PYTHON USING HOLT-WINTERS METHOD

S.K. Korsakova

Student of group 121701, BSUIR

N.V. Rusina

*Senior Lecturer, Department of
Economic Informatics, BSUIR*

V.M. Bandaryk

*Dean Faculty of Pre-University
Preparation and Occupational
Guidance of BSUIR,
PhD of Technical Sciences,
Associate Professor*

Annotation. Forecasting is always ahead of reality. It is one of the ways of planning, which allows to make preparations for situations that may arise in the future, thus helping to optimize the expenditure of resources: time, material, etc. Forecasting can be applied in any social sphere: in economics, education, medicine, business, production, etc. This article considers one of the methods of forecasting with the help of time series.

Keywords: time series, Holt-Winters method, forecasting.

УДК 004.4' 242:004.6-027.45

УСКОРЕННАЯ ОЦЕНКА НАДЁЖНОСТИ РАЗРАБАТЫВАЕМЫХ ПРИКЛАДНЫХ КОМПЬЮТЕРНЫХ ПРОГРАММ



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.К. Дик
Доцент кафедры электронной техники и технологии БГУИР,
кандидат технических наук,
доцент
sdick@bsuir.by



А.В. Будник
декан факультета инжиниринга 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; надёжность прикладных компьютерных программ; 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.К. Дик

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

А.В. Будник

Окончил Белорусский государственный университет информатики и радиоэлектроники. Основные направления научной деятельности – микроэлектроника, защита информации в технических системах.

Аннотация. Уточнены и систематизированы результаты исследований по оценке эксплуатационной надёжности разрабатываемых компьютерных программ, предназначенных для использования в различных сферах деятельности людей, в том числе в сфере *Big Data*.

Приводятся сведения по оценке ожидаемой начальной надёжности компьютерных программ и рекомендации по оценке продолжительности тестирования с целью обеспечения эксплуатационной надёжности планируемых к разработке прикладных компьютерных программ для сложных информационных систем.

Ключевые слова: прикладные компьютерные программы, прогнозируемая плотность ошибок, модели надёжности, тестирование.

Введение. В настоящее время проблема обеспечения надёжности разрабатываемых компьютерных программ стоит остро, поскольку по мнению ряда специалистов вклад программного обеспечения в ненадёжность сложных информационно-компьютерных систем может составлять 40 и более процентов [1].

Формы проявления ненадёжности прикладных компьютерных программ разнообразны: «зависание» программы, переполнение памяти, невозможность продолжения выполнения из-за некорректности входных данных, деление на нуль и т.д. Ненадёжность прикладных компьютерных программ большого размера (объёма) является следствием наличия в них скрытых неявных ошибок, оставшихся после выполнения процедуры её тестирования в течение отводимого времени на разработку программы. Специалисты по программированию пришли к выводу, что при оценке надёжности компьютерных программ в качестве единственно приемлемой количественной характеристикой измерения объёма программы является число исполняемых строк программного кода (в англоязычном варианте *LOC – Lines Of Code*) [2].

Прикладные компьютерные программы для сложных систем могут включать сотни тысяч и даже миллионы строк кода. Как пример, большой объём современного программного обеспечения иллюстрируется данными, приведёнными в таблице 1.

Таблица 1. Объём программного обеспечения

Программное средство (объект)	Космическая станция	Космический корабль	Boeing 777	Windows NT5	Windows 95	Windows XP
Количество строк кода, млн	40	10	7	35	5	40

Среднее число ошибок, приходящихся на тысячу строк кода (*KLOC*) варьируется от 5 до 50 ошибок [1]. Даже в системах, которые прошли строгий контроль качества (в англоязычном наименовании *Quality Assurance – QA*) все равно содержатся приблизительно 5 ошибок на 1000 строк кода [2]. Для ответственного программного обеспечения (ПО), к которому можно отнести операционные системы и другое системное ПО, к моменту поставки системы клиенту в нём может содержаться 0,04...0,15 ошибки на 1000 строк кода программы.

В прикладных компьютерных программах, которые протестированы только на предмет работоспособности функциональных возможностей, присутствует около 50 ошибок на 1000 выполняемых строк кода (рисунок 1).

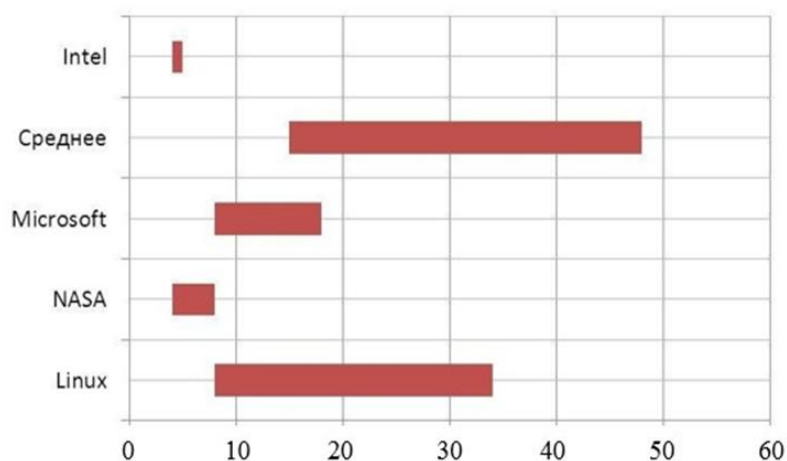


Рисунок 1. Среднее число ошибок, приходящихся на 1000 строк кода,

для прикладных программных средств, прошедших тестирование

По данным [4] отношение числа ошибок, не выявленных при проектировании прикладной компьютерной программы, к общему числу команд программы (обозначим общее число команд программы через B) лежит в диапазоне от 0,25 до 10 на 1000 команд.

Примерное значение числа команд B можно определить как [6]

$$B = E_L L, \quad (1)$$

где E_L – коэффициент расширения кода (увеличения числа команд компьютерной программы относительно числа выполняемых строк кода L).

Согласно работе [6] можно считать, что одна строка кода компьютерной программы транслируется в 10 машинных команд, т. е. $E_L = 10$.

Актуальность. В соответствии с ГОСТ 27.205-1997 [3] о прогнозной надёжности разрабатываемой компьютерной программы будем судить по интенсивности проявления ошибок (λ), обусловленных дефектами её проектирования. Поскольку проявление ошибки при использовании программы приводит к невыполнению задачи пользователя, что равносильно отказу программы, то далее будем употреблять термин «интенсивность отказов программы». Использование λ в качестве характеристики надёжности программных средств удобно тем, что её можно оценивать по прогнозируемой плотности ошибок F в компьютерной программе [5]. Плотность ошибок F это отношение вида

$$F = \frac{\text{Количество возможных ошибок в компьютерной программе}}{\text{Число строк кода компьютерной программы}}.$$

Прогнозирование плотности отказов не требует знания среды исполнения компьютерной программы и, следовательно, подходит для ранних этапов разработки программного обеспечения. Когда информация о предполагаемой среде исполнения становится доступной, прогнозируемая плотность отказов может быть преобразована в прогнозируемую интенсивность отказов. Актуальными являются получение прогнозного значения начальной плотности ошибок, т.е. плотности ошибок компьютерной программы после написания её кода и устранения нарушений правил языка программирования (обозначим через F_0) и алгоритма преобразования её в начальную интенсивность отказов λ_0 разрабатываемой компьютерной программы.

Прогнозное значение начальной плотности ошибок. Определение прогнозного значения начальной плотности ошибок F_0 рассматривалось в [7–10]. На основе анализа установлено, что возможным является использование модели RL-92-52 [5], согласно которой F_0 можно представить в виде произведения

$$F_0 = F_B \cdot D \cdot S = F_B \cdot \prod_{i=1}^m D_i \cdot \prod_{j=1}^n S_j, \quad (2)$$

где F_B – базовая (средняя) плотность ошибок для компьютерных программ, используемых в данной отрасли, области применения; D – метрика, учитывающая производственную среду разработки программного обеспечения, в том числе особенность организации, разрабатывающей компьютерную программу, характеристику группы программистов, их квалификацию; S – метрика, учитывающая характеристики разрабатываемой компьютерной программы; D_i – поправочный коэффициент, учитывающий влияние на значение F_0 i -го фактора, описывающего производственную среду разработки компьютерной программы; S_j – поправочный коэффициент, учитывающий j -ю

характеристику компьютерной программы; m – число факторов, принимаемых во внимание при описании метрики D ; n – число характеристик компьютерной программы, выбранных для описания метрики S .

Коэффициенты метрик D и S могут принимать значения как меньше единицы, так и больше неё в зависимости от того, как фактор влияет на плотность ошибок.

Метрику D модели (2) в соответствии с [8] предлагается определять в виде произведения двух коэффициентов:

$$D = D_{\text{орг}} \cdot D_{\text{квал}},$$

где $D_{\text{орг}}$ – коэффициент, характеризующий особенность организации, разрабатывающей прикладную компьютерную программу; $D_{\text{квал}}$ – коэффициент, учитывающий квалификацию и опыт программистов.

Вопросы выбора коэффициента $D_{\text{орг}}$ рассматриваются в [5]. Для выбора коэффициента $D_{\text{квал}}$ можно использовать рекомендации работы [10] (таблица 1).

Таблица 1. Значения коэффициента $K_{\text{квал}}$

Квалификация и опыт программиста	Значение $K_{\text{квал}}$
1 Студент, освоивший программирование на уровне программы учебной дисциплины высшего технического учебного заведения	2,0
2 Младший программист (<i>Junior Developer</i>)	1,3
3 Программист (<i>Middle Developer</i>)	1,0
4 Ведущий программист (<i>Senior Developer</i>)	0,7

Выбор коэффициентов S_j , принимаемых во внимание для компьютерной программы, удобно сделать в соответствии с документом «Укрупнённые нормы затрат труда на разработку программного обеспечения. Постановление Министерства труда и социальной защиты Республики Беларусь 27.06.2007 № 91». Этот документ позволяет определить нормы времени на разработку компьютерных программ с учётом основных их характеристик. С учётом документа [11] рекомендуется принять во внимание коэффициенты S_j , учитывающие следующие характеристики разрабатываемой компьютерной программы:

S_1 – категорию сложности компьютерной программы;

S_2 – используемые средства разработки компьютерной программы (язык программирования; операционная система компьютера; сеть функционирования: локальная, глобальная; другие факторы);

S_3 – степень новизны разрабатываемой компьютерной программы;

S_4 – степень использования стандартных модулей в компьютерной программе.

Для выбора значений коэффициентов $S_1 \dots S_4$ рекомендуется использовать данные документа [11] с учётом гипотезы о том, что число возможных ошибок в компьютерной программе, следовательно, и плотность ошибок, прямо пропорционально затратам труда на её разработку.

Преобразование плотности ошибок в интенсивность отказов программы. В работе [10] на основе экспериментальных данных о надёжности прикладных компьютерных программ [5], показано, как по прогнозному значению начальной плотности ошибок F_0 , объёму программы в исполняемых строках программного кода L ,

области применения программы, а также быстродействию R процессора компьютера, получить ожидаемый (примерный) уровень начальной интенсивности отказов λ_0 :

$$\lambda_0^{(l)} = (2520 C^{(l)} V_{\text{пик}} N_0^{(l)}) = 60 K_{\Sigma}^{(l)} V_{\text{пик}} F_0^{(l)} L 10^{-6}, \text{ч}^{-1}, \quad (3)$$

где верхний индекс l означает, что соответствующие параметры относятся к прикладной компьютерной программе l -й области применения; F_0 – плотность ошибок компьютерной программы после написания её кода и устранения нарушений правил языка программирования, размерность: ошибка / строка кода; $C^{(l)}$ – коэффициент проявления ошибок программы, размерность: 1/ошибка; $V_{\text{пик}}$ – пиковая скорость выполнения команд компьютерной программы, представляющая собой возможное число прогонов компьютерной программы в течение одного часа для пикового быстродействия процессора; N_0 – начальное число ошибок (до выполнения тестирования); $K_{\Sigma}^{(l)}$ – суммарный коэффициент увеличения интенсивности отказов, обусловленный совместным действием изменчивости входных данных и рабочей нагрузки на компьютерную программу; L – прогнозное значение числа выполняемых строк программного кода (LOC).

В случае затруднений в определении прогнозного значения L разрабатываемой компьютерной программы рекомендуется обратиться к документу [11].

Коэффициент $C^{(l)}$ модели (3) показывает, как значение начальных ошибок $N_0^{(l)}$ в прикладной компьютерной программе, предназначенной для её применения в l -й области, трансформируется в начальную интенсивность отказов λ_0 при пиковой скорости $V_{\text{пик}}$ выполнения команд компьютерной программы, определяемой как

$$V_{\text{пик}} = \frac{R_{\text{пик}}}{B}, \quad (4)$$

где $R_{\text{пик}}$ – пиковое значение быстродействия процессора компьютера, указанное в технической документации, размерность: «операций в секунду»; B – прогнозное число команд (операций), которое будет выполняться программой в течение одного прогона для экстремального набора входных данных и экстремальной рабочей нагрузки на компьютерную программу со стороны эксплуатационной среды.

В первом произведении модели (3) коэффициент (число) 2520 получен с учётом того, что рассматривается скорость $V_{\text{пик}}$ выполнения команд компьютерной программы в час, а реальное быстродействие R процессора принято, как $R = 0,7 \cdot R_{\text{пик}}$.

Достоверность начальной интенсивности отказов компьютерной программы λ_0 во многом определяется точностью оценки значения B . Использование равенства (1) для получения значения B даёт слишком ориентировочную оценку. Рекомендуется делать экспертную оценку числа команд B , выполняемых прикладной компьютерной программой за один прогон для случая экстремального режима входных данных и экстремальной нагрузки 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.). Этап экспертной оценки B является достаточно сложным.

Используя экспериментальные данные о надёжности прикладных компьютерных программ, в [10] получены значения коэффициентов $C^{(l)}$ и $K_{\Sigma}^{(l)}$ для разных областей применения, для которых разрабатываются прикладные компьютерные программы (таблица 2). Эти коэффициенты могут рассматриваться в качестве справочных, используемых для оценки надёжности разрабатываемых прикладных компьютерных программ.

В таблицу 2 также включены справочные значения базовых плотностей ошибок, полученные на основе исследования надёжности прикладных компьютерных программ разных областей (сфер) их применения [5].

Таблица 2. Значения коэффициентов для компьютерных программ различных областей применения

Область применения компьютерной программы	Коэффициент $C, \times E-07, 1/\text{ошибка}$	Коэффициент K_{Σ}	Значение $F_{\text{б}}, \text{ошибка} / \text{строка кода}$	Средний процент времени прогона программы при её тестировании в течение календарного времени $T_{\text{календ}}$
1 Авиация, космос	1,246	5,23	0,0128	8
2 Мониторинг и обеспечение безопасности	0,238	1,00	0,0092	43
3 Телекоммуникации, мобильные устройства	2,738	11,5	0,0078	3,5
4 Управление производственными процессами	0,754	3,17	0,0018	14
5 Автоматизированные системы управления, технологии <i>Big Data</i>	4,563	19,2	0,0085	2,5
6 Разработка программ, моделирование, обучение	3,353	14,1	0,0123	3
Среднее	2,103	8,83	0,0087	12

Значение λ_0 , полученное по модели (3), обычно не отвечает требованиям заказчика, поэтому выполняют тестирование компьютерной программы с целью поиска и устранения скрытых ошибок, которые гипотетически могут иногда проявляться в зависимости от набора входных данных и уровня нагрузки программы со стороны эксплуатационной среды. Используя работу [12], требуемое процессорное время тестирования можно определить по формуле

$$T_{\text{календ}} = \frac{B \ln(Q)}{60 K_{\Sigma}^{(l)} R r^{(l)}} 10^{-4} = \frac{L K_B \ln(Q)}{60 K_{\Sigma}^{(l)} R r^{(l)}} 10^{-4}, \text{ ч}, \quad (5)$$

где Q – коэффициент тестирования; $r^{(l)}$ – средний процент времени выполнения (прогона) компьютерной программы при её тестировании в течение рабочей календарной продолжительности для прикладных компьютерных программ l -й области применения (см. таблицу 2); K_B – коэффициент увеличения числа команд компьютерной программы относительно числа выполняемых строк программного кода L .

Коэффициент K_B учитывает, как язык программирования, так и специфику процедур, выполняемых компьютерной программой.

В формулу (5) пиковое быстродействие процессора R следует подставлять в размерности «операций в секунду», показатель L – в строках кода (LOC).

Значение Q находят как

$$Q = \frac{\lambda_0}{\lambda_{\text{экс}}}, \quad (6)$$

где $\lambda_{\text{экс}}$ – допустимая эксплуатационная интенсивность отказов компьютерной программы (интенсивность проявления ошибок).

В качестве основной характеристики безотказности выполнения программой своих функций, согласно [4], 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 [3], эту вероятность (обозначим через P_1) можно определить, как

$$P_1 = \exp(-\lambda_{\text{экс}} t_1), \quad (7)$$

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 (7) в качестве времени t_1 следует рассматривать процессорное время обработки данных (без учёта времени ожидания).

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 это среднее число сеансов через η). Тогда вероятность того, что оставшиеся ошибки в прикладной компьютерной программе не проявятся в течение заданного календарного времени τ , можно определить, используя выражение

$$P(\tau) = (P_1)^\eta \tau. \quad (8)$$

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

Список литературы

- [1] 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/files/5357ec/dd0cd8/50af39/000000/seminar_metody_obespecheniya_apparatno-programmnoy_nadezhnosti_vychislitelnyh_sistem.pdf. – Дата доступа: 31.01.2024.
- [2] Программное обеспечение – источник всех проблем. [Электронный ресурс]. – 2018. – Режим доступа: <http://www.williamspublishing.com/PDF/5-8459-0785-3/part1.pdf>. – Дата доступа: 31.01.2024.
- [3] Шубинский, И. Б. Функциональная надёжность информационных систем. Методы анализа / И. Б. Шубинский. – М. : «Журнал Надёжность», 2012. – 296 с.
- [4] 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.
- [5] McCall, J. A. Software reliability, measurement, and testing guidebook for software reliability measurement and testing / J. A. McCall [et al.] // Final Technical Report, Vol II (of two) ; RL-TR-92-52. – Science Applications International Corp., Research Triangle Institute : Rome Laboratory, NY 13441-5700, 1992. – 256 p.

- [6] Чуканов, В. О. Надёжность программного обеспечения и аппаратных средств систем передачи данных атомных электростанций : учеб. пособие / В. О. Чуканов. – М. : МИФИ, 2008. – 168 с.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] Оценка ожидаемой надёжности прикладных программных средств для компьютерных информационных систем / С. М. Боровиков [и др.] // Информатика. – 2021. – Т. 18, № 1. – С. 84-95. <https://doi.org/10.37661/1816-0301-2021-18-1-84-95>
- [11] 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-ukрупnennyh-norm-zatrat-truda>. – Дата доступа: 31.01.2024.
- [12] 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.

Авторский вклад

Лэ Ван Там – разработка моделей по оценке ожидаемой надёжности прикладных компьютерных программ разных областей применения

Боровиков Сергей Максимович – общее руководство исследованиями и разработка моделей по оценке надёжности прикладных компьютерных программ.

Дик Сергей Константинович – постановка задачи и разработка модели определения прогнозного значения начальной плотности ошибок в прикладной компьютерной программе.

Будник Артур Владимирович – постановка задачи по обеспечению эксплуатационной надёжности прикладных компьютерных программ, подготовка справочных параметров для оценки надёжности прикладных компьютерных программ для разных областей применения.

ACCELERATED ASSESSMENT OF THE RELIABILITY OF DEVELOPED APPLICATION SOFTWARE

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

S.M. Borovikov
*Associate Professor,
Department of
Information Computer
Systems Design, PhD of
Technical sciences,
Associate Professor*

S.K. Dick
*Associate Professor,
Department of
Electronic Engineering
and Technology, PhD of
Technical sciences,
Associate Professor*

A.V. Budnik
*Dean of the Faculty of
Engineering and
Technology of the
BSACE, PhD of
Technical sciences,
Associate Professor*

Abstract. The results of research to assess the operational reliability of the software being developed, intended for use in various fields of human activity, including in the field of Big Data, have been clarified and systematized.

Information is provided on assessing the expected initial reliability of software and recommendations on assessing the duration of testing in order to ensure the operational reliability of application software planned for development for complex information systems.

Keywords: application software, predicted density fault, reliability models, testing.

УДК 004.042

DEVELOPMENT OF A METHOD FOR USING AUTOENCODER TO SEARCH FOR ANOMALIES IN CLOUD DATA



C.S. Dzik



I.I. Piletskii



T.A. Asipovich

C.S. Dzik

Graduate student BSUIR, Utech Solutions, software and data engineer, conduct scientific research of anomaly detection using autoencoder artificial neural network

I.I. Piletski

PhD, Associate Professor of the department of Informatics Department of Belarusian State University of Informatics and Radioelectronics. In the field of IT for over 50 years. Participation in the development of several dozen large projects: chief designer of the project, chief architect of software and information support, project manager, head of department.

T.A. Asipovich

Graduated from the Belarusian State Economic University. The area of scientific interests is related to the study of problems of e-commerce and entrepreneurship in the field of information technology.

Annotation. A methodology for conducting an experiment to search for anomalies in a cloud data array using an autoencoder is proposed. The technique was developed using an example and for use in analyzing the telemetry results of a solar power plant to search for defects and anomalies in its operation. However, with minor adaptation, it can be used to search for anomalies in other types of cloud data.

Keywords: autoencoder, PV Modules, cloud data, anomaly search, data defects.

Introduction. According to data quality experts, data is of high quality when it satisfies the requirements of its intended use. In other words, companies know that they have good quality data when they are able to use it to communicate effectively with their constituents, determine clients' needs, and find effective ways to serve their client base [1-4].

This data quality definition is broad enough to help companies with varying products, markets, and missions to understand if their data is up to standards. Data quality is not good or bad, high or low. It is a range or an indicator of operability of the data that passes through a company. Data quality management ensures the context-dependent process of improvement of suitability of the data, which is used for analysis and decision-making. The goal is to provide the vision of the «health» of the data by applying different processes and technologies to the increasingly complex data sets [1-4].

We want to be sure that when we take advantage of the cloud to help data managing, we define data quality parameters at the same time. The most obvious and compelling way to achieve the goal is to make sure we perform automatic data quality checks for all our data, wherever they are - in the cloud or elsewhere. We must always perform an on-site data quality check.

Virtually every company that works with data has a certain data quality (DQ) monitoring system. Some companies even hire an entire department that deals with the issue. This option is very expensive. In addition, most data quality checks are hard-coded and rule-based. In the event of a failure, the system notifies you of the risk indicator. Such rules are often critical to business continuity. For example, we cannot have a missing customer ID or a «risk profile» variable with an incorrect value. As the amount of data grows, you cannot specify a rule for work with each attribute; not to mention the difficulty of working with hard-coded multidimensional control checks.

The best option is automated DQ (data quality) checks using Machine Learning to detect anomalies that we don't even need to explicitly program.

For the DAD task the normal data easily reconstructed by the autoencoder, while the anomalous object for the model will be difficult to reconstruct. An autoencoder is a type of artificial neural network used to learn efficient data codings in an unsupervised manner.

Autoencoder are can: accept an input set of data; internally compress the data into a latent-space (low dimensional) representation; reconstruct the input data from the latent representation.

To accomplish this task, an autoencoder uses two components: an encoder and a decoder.

The encoder accepts the input data and compresses it into the latent-space representation. The decoder then attempts to reconstruct the input data from the latent space [5].

The aim of an autoencoder is to learn a representation (encoding) for a set of data, typically for dimensionality reduction. Along with the reduction side, a reconstructing side is learnt, where the autoencoder tries to generate from the reduced encoding a representation as close as possible to its original input [5 – 9].

The purpose of this work is to develop an experimental methodology for analyzing a large array of data using the example of telemetry results of solar power plants.

Development of the methodology. Figure 1 shows the general diagram of an autoencoder-type neural network. When analyzing big data, the autoencoder input for training must receive input data vectors in the form of parameter values that do not have defects or anomalies. That is, average statistical data that will depend on the problem being solved. After internal transformations in the autoencoder, its output must contain vectors of output data identical to the input. After training, when data vectors with defects or anomalies are fed to the input of the autoencoder, vectors without defects will be obtained at its output. By comparing the input and output vectors with the data, it will be possible to draw a conclusion about the presence of anomalous parameters in the input data vector.

Thus, to conduct an experiment to search for anomalies in the telemetry results of a solar power plant, it is necessary to develop an internal model of the autoencoder, develop a structure of input and output vectors for training and operation of the autoencoder, prepare a dataset for training the autoencoder, develop a mechanism and software for comparing the result of the autoencoder and conduct experiment on empirical data.

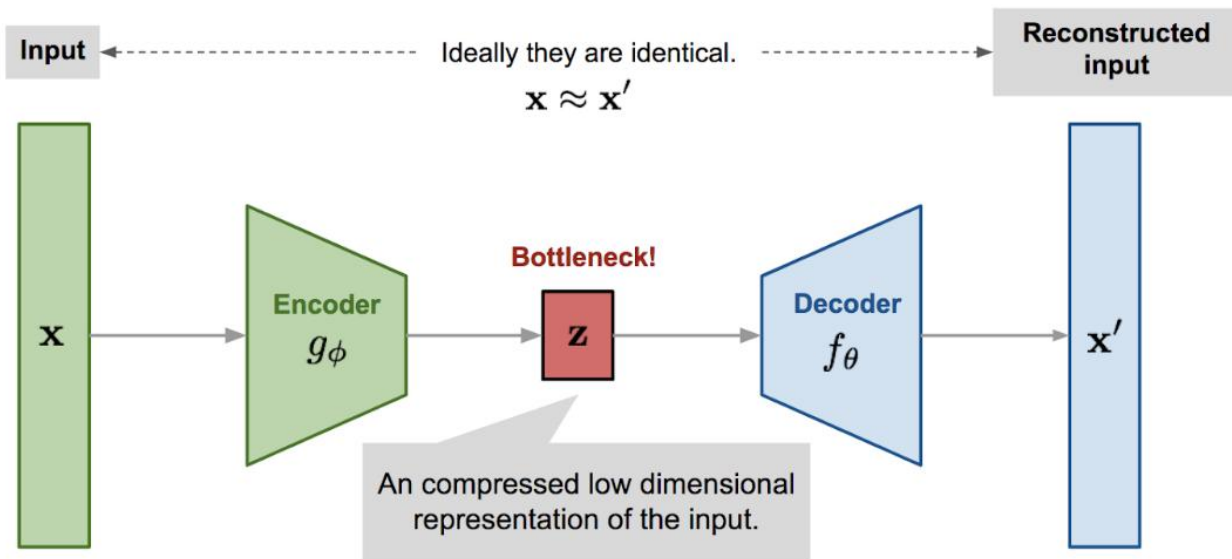


Figure 1. Example autoencoder neural network

Taking into account the specifics of the telemetry results of the operating parameters of the PV Modules of the power plant, we will use a vector of dimension 200×5 as a unit vector of input data, which will include the telemetry results from 10.00 to 18.00 hours of one day of the following parameters: timestamp, voltage, current, temperature in the PV Module housing, the level of illumination. The test data vector had a dimension of 200×3 and included the following parameters: timestamp, temperature in the PV Module housing, illumination level. At the output of the autoencoder, in both cases, a vector of 200×2 was obtained, that is, when test data is supplied to the input, the autoencoder must, based on the training results, restore the voltage and current values based on the temperature in the PV Module housing and the illumination level. In addition, to train the neural network, telemetry data of stably operating PV Modules (without defects and reduced efficiency) were used, selected as a result of direct analysis and the use of other methods of searching for anomalies in the operation of PV Modules.

To exclude telemetry results taken on days with partly cloudy and cloudy days from the training dataset, you need to use a filter that will leave data only on clear sunny days with low clouds. The filtering parameters are selected empirically and have the following values: current – $0 - 15$ A, illumination – $360 - 1500$ W/m², derivative with respect to current – $-0.2 - 0.2$, derivative with respect to illumination – $-3.8 - 3, 8$.

Figure 2 shows the voltage curves on Module 2.3_10 (defective) and Module 2.3_9 (good) during a sunny day. It can be seen that the voltage curves for these PV Modules are different. This is due to a malfunction in the Module 2.3_10: after the inverter entered the energy removal mode, the protective diode tripped, which disconnected a third of the cells from the circuit, so during the day the voltage on this PV Module is reduced.

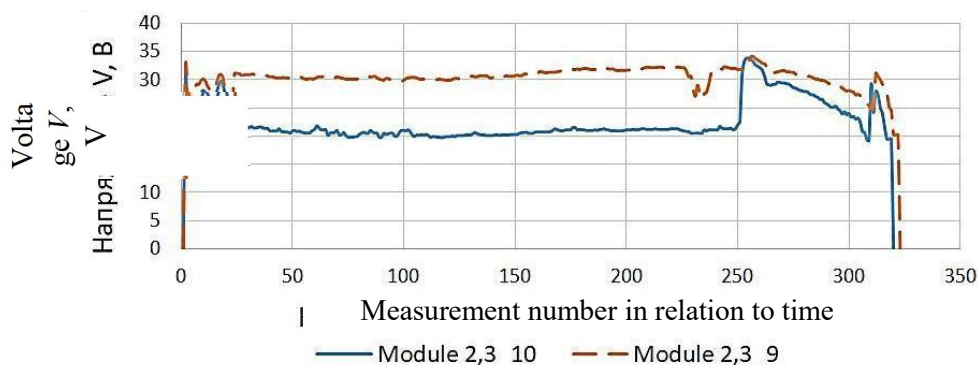


Figure 2. Change of voltage in Module 2,3_9 and Module 2,3_10 depending on time

Figure 3 shows the current curves on Module 2.3_1 (defective) and Module 1.5_12 (good) during a sunny day. The difference in the curves is due to the fact that in the cloud storage some of the points for the Module 2.3_1 PV Module are missing, and therefore it is shifted to the left. At the same time, the current curve for the PV Module Module 1.5_12 shows that at the beginning of the day there was cloudiness, which caused the telemetry collection device for the chain of PV Modules String 2.3 to be turned off. A defect in the data is the absence of some points on the current curve. Otherwise, the Module 2.3_1 works fine. The proposed filter will eliminate such situations due to the fact that points taken for the study will be taken during the absence of sharp changes in the readings of current strength and illumination.

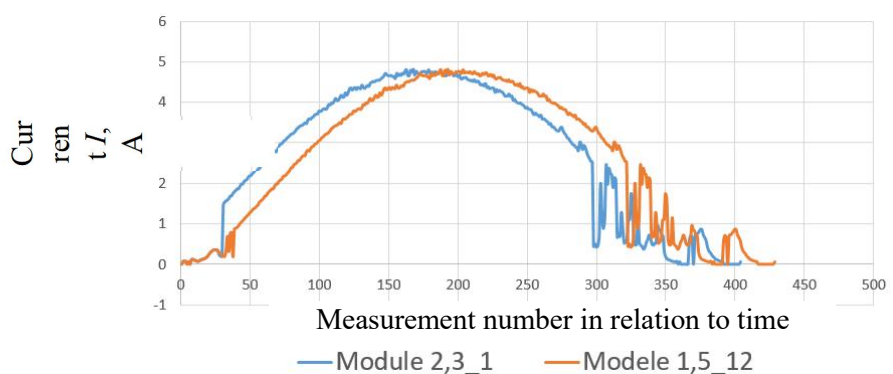


Figure 3. Change of current in Module 2,3_1 and Module 1,5_12 depending on time

The autoencoder developed to solve the problem consists of seven layers of neurons: input, output and five internal layers.

As a result of the autoencoder operation, current and voltage curves will be obtained for each PV Module depending on the illumination and temperature in the housing. These curves will be subject to comparison with curves obtained as a result of telemetry collection in automatic mode. A list of PV Modules with a curve divergence of more than 5% will be saved in a special file.

Conclusion. An experimental methodology, an autoencoder model, and a software algorithm for searching for anomalies in cloud data using the example of the results of collecting telemetry from solar power plants have been developed. The proposed approach will allow not only to search for anomalies in the operation of a solar power plant, but also to use it to search for anomalies in cloud data of a different nature and structure.

Reference list

- [1] ИСО 8000-2 Качество данных. Часть 2. Словарь (ISO 8000-2, Data quality - Part 2: Vocabulary)

[2] ИСО/ТС 8000-110 Качество данных. Часть 110. Основные данные. Обмен данными характеристик. Синтаксис, семантическое кодирование и соответствие спецификации данных (ISO 8000-110, Data quality - Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification)

[3] ИСО/ТС 8000-120 Качество данных. Часть 120. Основные данные. Обмен данными характеристик. Происхождение (ISO/TS 8000-120:2009, Data quality - Part 120: Master data: Exchange of characteristic data: Provenance)

[4] Data quality [\[\[Online Resource\]](https://en.wikipedia.org/wiki/Data_quality) – Access Mode: https://en.wikipedia.org/wiki/Data_quality // Access Date: 14.02.2022

[5] U.A. Vishniakou, O.S. Koval, M.G. Mozdurani Shiraz. Use of Neural Networks for Detection and Recognition of the Anomalies in Enterprise Corporate Information System. Doklady BGUIR, 86 (4), 86-92.

[6] Alexander Prosak, Amitava Gangopadhyay and Hemant Garg A New Machine Learning Approach for Anomaly Detection Using Metadata for Model Training. EasyChair Preprint No 829 <https://easychair.org/publications/preprint/Sf34>

[7] Zhou, Chong, and Randy C. Paffenroth. "Anomaly detection with robust deep autoencoders." Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017.

[8] Di Mattia, Federico, et al. "A Survey on GANs for Anomaly Detection." arXiv preprint arXiv:1906.11632 (2019).

[9] Malhotra, Pankaj, et al. "LSTM-based encoder-decoder for multi-sensor anomaly detection." arXiv preprint arXiv:1607.00148 (2016).

Авторский вклад

Константин Сергеевич Дик – разработка архитектур автоэнкодера, программного обеспечения фильтрации и сравнения данных телеметрии, оформление результатов работы.

Иван Иванович Пилецкий – постановка задачи исследования, проработка плана работ, анализ полученных результатов.

Татьяна Анатольевна Осипович – анализ результатов телеметрии, разработка способов фильтрации данных телеметрии и сравнения их с результатами работы автоэнкодера.

РАЗРАБОТКА МЕТОДА ИСПОЛЬЗОВАНИЯ АВТОЭНКОДЕРА ДЛЯ ПОИСКА АНОМАЛИЙ В ОБЛАЧНЫХ ДАННЫХ

Дик К.С.

Аспирант БГУИР, Ютех
Солюшнс, инженер по
программному обеспечению и
данным

И.И. Пилецкий

к.ф.-м.н., доцент кафедры
информатики Белорусского
государственного университета
информатики и
радиоэлектроники. В сфере IT
более 50 лет. Участие в
разработке нескольких
десятков крупных проектов

Т.А. Осипович

Доцент кафедры экономики
БГУИР, кандидат
экономических наук

Аннотация. Предложена методика проведения эксперимента по поиску аномалий в массиве облачных данных с использованием автоэнкодера. Методика разработана на примере и для использования при анализе результатов телеметрии солнечной электростанции на предмет поиска дефектов и аномалий в её работе. Однако, при незначительной адаптации, может быть использована при поиске аномалий в других видах облачных данных.

Ключевые слова: автоэнкодер, солнечные панели, облачные данные, поиск аномалий, дефекты данных.

УДК 612.087.1

ПРОГРАММНЫЙ МОДУЛЬ ОБРАБОТКИ ШУМОПОДОБНЫХ БИМЕДИЦИНСКИХ СИГНАЛОВ НА ОСНОВЕ КОРРЕЛОГРАММЫ



Малицкий В.А.
Младший научный
сотрудник НИЧ
БГУИР,
магистрант,
vsevolod.malitskiy@
gmail.com



Осипов А.Н.
Заведующий
лабораторией НИЧ
БГУИР,
доцент, к.т.н.,
osipov@bsuir.by



Клюев А.П.
Старший
преподаватель
кафедры ИПиЭ
БГУИР
kluev@bsuir.by



Тяньбо Ма
Аспирант БГУИР
714922741@qq.com

В.А.Малицкий

Работает в Белорусском государственном университете информатики и радиоэлектроники. Область научных интересов – цифровая обработка биомедицинских сигналов.

А.Н.Осипов

Работает в Белорусском государственном университете информатики и радиоэлектроники. Область научных интересов – биомедицинская инженерия.

А.П.Клюев

Работает в Белорусском государственном университете информатики и радиоэлектроники. Область научных интересов – системы с биологической обратной связью.

Тяньбо Ма

Учится в Белорусском государственном университете информатики и радиоэлектроники. Область научных интересов – биомедицинская инженерия.

Аннотация. В данной статье приведено описание и алгоритм работы программного модуля на основе метода корреляционного анализа, разработанного при помощи пакета прикладных программ *MatLab* для исследования шумоподобных биомедицинских сигналов.

Ключевые слова. Корреляционный анализ, коррелограмма, коэффициент корреляции, массив данных.

Введение. Для расчета значимых параметров биосигналов традиционно используют широкий спектр методов: от статистического и анализа гистограмм распределения до спектрального анализа и нелинейных методов [1]. Для анализа шумоподобных сигналов традиционно используется частотно-временное преобразование [2]. В данной статье для этих целей предлагается использовать метод, основанный на вычислении коррелограмм.

Программный модуль. Программный модуль, созданный с помощью пакета прикладных программ *MatLab*, осуществляет анализ данных последовательностей двух сигналов на основе корреляционного анализа. Корреляционный анализ – статистический метод изучения взаимосвязи между двумя и более последовательностями [3]. Программный модуль осуществляет расчёт коэффициентов корреляции в зависимости от

размера выборки (окна), а затем, по вычисленным данным, выполняются построение коррелограммы. Данный процесс осуществляется в несколько этапов.

На первом этапе загружаются последовательности исследуемых сигналов, записанных в формате *.txt* файлов, в среду *MatLab* для их обработки. Благодаря встроенной функции данные из *.txt* файлов записываются в отдельные массивы. Сами массивы состоят из одинакового кол-ва элементов N . При различных длинах сигналов пользователю следует сократить более длинный массив до размера меньшего массива, либо дополнить меньший массив нулями до размеров большего. В программе предусмотрена возможность смещения данных в массивах, благодаря которой выбирается наиболее интересующий участок сигналов для исследования.

На втором этапе задаётся размер окна M (количество отсчётов), который влияет на разрешающую способность. Выборка окон производится последовательно без перекрытия и не включают в себя элементы предыдущего окна. Размер окна и их количество K определяются в соответствии с выражением $K = N/M$.

Следует отметить, что число окон K должно быть целым. Следовательно, в ситуациях получения нецелого значения K изменяется размер окна, дополняются или сокращаются массивы. В предлагаемой программе избыточные элементы исключаются и далее не используются. Это целесообразно до момента, пока количество отбрасываемых элементов не превышает половину размера окна.

На третьем этапе выполняется расчёт коэффициентов корреляции в каждом окне и запись данных в двумерный массив. Общая формула вычисления коэффициентов корреляции $r_{12}(n)$ двух исследуемых последовательностей сигналов $x_1(n)$ и $x_2(n)$, содержащих по N элементов, определяется как:

$$r_{12}(j) = \frac{1}{N} \sum_{n=0}^{N-1} x_1(n)x_2(n+j), \quad (1)$$

где j – величина сдвига или интервал, на который $x_2(n)$ смещается относительно $x_1(n)$.

В рассматриваемом случае $j = 1, 2, \dots, M$ относительно первого элемента $x_1(n)$, а общее кол-во сдвигов $x_2(n)$ в окне равно размеру окна. Сдвиг выполняется внутри окна и не затрагивает значения вне этого окна. Последнее значение в этом промежутке на следующей итерации становится первым в этой последовательности.

Рассчитанные коэффициенты корреляции записываются в двумерный массив. В данном двумерном массиве кол-во строк соответствует размеру окна, а количество столбцов – количеству окон. Данный формат удобен для последующей обработки специальными функциями *MatLab*. В программе имеется возможность изменения размерности массива и обработки его не как матрицы, а как строки с последовательной записью результатов расчёта.

На четвертом этапе, применяя функцию *pcolor*, выполняется построение графика коррелограммы. Вдоль оси X откладываются номера окон, вдоль оси Y – сдвиг в окнах, а значение корреляционного коэффициента отображается насыщенностью цвета.

Стоит упомянуть о нормировании сигнала. При использовании функции *pcolor* и других встроенных функциях максимальная и минимальная интенсивность цвета определяется по максимальному и минимальному значению коэффициентов корреляции автоматически. Она может быть необходима для более удобной оценки количественных данных человеком.

Заключение. Таким образом в данной статье приводится описание разработанного авторами программного модуля вычисления коррелограммы для обработки шумоподобных биомедицинских сигналов. Поэтапно описано функционирование данного

модуля, с необходимыми пояснениями его особенностей. Результаты работы могут быть использованы при проектировании медицинской диагностической аппаратуры.

Список литературы

[1] Биомедицинские сигналы и изображения в цифровом здравоохранении: хранение, обработка и анализ: учебное пособие / В.С. Кубланов, А.Ю. Долганов, В.Б. Костоусов [и др.]; [под общ. ред. В. С. Кубланова]; Мин-во науки и высш. образования РФ. — Екатеринбург: Изд-во Урал. ун-та, 2020.— 240 с.

[2] Меженная, М. М. Аппаратно-программные средства электростимуляции опорно-двигательного аппарата человека на основе частотно-временного анализа биоэлектрической активности мышц: автореф. дисс. ... кандидата технических наук: 05.11.17 / Меженная М. М.; науч. рук. А. Н. Осипов. - Минск: БГУИР, 2012. - 22 с

[3] Корреляционный анализ [Электронный ресурс] / Режим доступа - <https://www.statmethods.ru/statistics-metody/korrelyatsionnyj-analiz/> - 25.01.2024

Авторский вклад

Авторы внесли равноценный вклад.

A SOFTWARE MODULE FOR PROCESSING NOISE-LIKE BIOMEDICAL SIGNALS BASED ON A CORRELOGRAM

Malitskiy V.A.
*Junior Researcher of
the BSUIR Research
Institute, master's
student,*

Osipov A.N.
*Head of the laboratory
of the BSUIR Research
Institute, Associate
Professor, PhD.,*

Klyuev A.P.
*Senior lecturer of the
Department of IPiE
BSUIR*

Tianbao Ma
*Graduate student of
BSUIR*

Annotation. This article describes and describes the algorithm of the software module based on the correlation analysis method developed using the MatLab application software package for the study of noise-like biomedical signals.

Keywords. Correlation analysis, correlogram, correlation coefficient, data array.

УДК 004.62:338.012

МОДЕЛИРОВАНИЕ И ФОРМИРОВАНИЕ ПРОСТРАНСТВЕННО-ВРЕМЕННОГО ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ПРИ ИНФОРМАТИЗАЦИИ, НА ПРИМЕРЕ ОТРАСЛИ МЕЛИОРАЦИИ



Вахонин Н.К.

*Профессор кафедры
компьютерных информационных
технологий МИТСО, кандидат
технических наук, доцент
Nik.vahonin@mail.ru*

Н.К. Вахонин

Окончил Белорусскую государственную сельскохозяйственную академию, Московский физико-технический институт. Область научных интересов связана с системным анализом, моделированием, разработкой систем принятия решений и их информационного обеспечения, на основе мониторинга и формирования пространственно-временных БД больших природно-технических систем в отрасли сельскохозяйственная мелиорация.

Аннотация. В статье приведены результаты системного анализа проблемы разработки системы поддержки принятия решений в больших, пространственно-распределенных природно-технических системах, со стохастическими входными воздействиями, на примере информатизации мелиоративной отрасли Выделена единая, связанная общими экономико-экологическими целями и неразрывным процессом динамики воды, система мелиоративный сельскохозяйственный объект, подлежащий оптимизации на всех этапах своего жизненного цикла (проектирование, создание, использование, контроль реконструкция). Сформулированы альтернативные варианты моделей функционирования системы и их информационного обеспечения. Показано, что выбор их оптимума сам является задачей целе ориентированного принятия решений, предопределяющего оценку эффективности информатизации, по результатам ее использования в реальном секторе экономики. Представлен ряд информационных средств, разработанных для мелиорации.

Ключевые слова: системный анализ, предметная область, реальный сектор экономики, большие системы, материальные системы, мелиоративная отрасль, мелиоративный сельскохозяйственный объект, информатизация, цифровая экономика, система поддержки принятия решений, оптимизация, цели, критерии, экономическая эффективность, этапы жизненного цикла систем, проектирование информационных систем, диаграммная символика, модель функционирования системы, уравнения с распределенными параметрами модель черный ящик, нейросетевая модель, стохастические входные воздействия, пространственно-временное информационное обеспечение, геореляционные базы данных, большие данные, вращение данных, геоинформационные системы.

Введение. В реализуемой в Беларуси Государственной программе Цифровое развитие Беларуси [1], наряду с быстрым формированием новой, инвестиционно привлекательной отрасли «чисто» информационно-коммуникационных технологий: электронная торговля, мобильные приложения, электронные игры, интернет-технологии передачи информации, выполнение заказов на оказание государственных услуг и т.п.,

особую актуальность имеет, и рассматривается как основа развития экономики страны информатизация материальных технологий и систем всех традиционных отраслей (предметных областей), реального сектора экономики (энергетика, промышленность, сельское хозяйство, транспорт, строительство и т.п.). –цифровая экономики.

При этом под повсеместно используемым на бытовом уровне словосочетании цифровизация экономики следует понимать не перевод на компьютерную основу имеющихся докомпьютерных, по неизбежности являющихся предельно упрощенными и, соответственно, неточными, методов принятия решений, базирующихся на экспертно-нормативных соображениях, а реинжиниринг технологий принятия решений с осуществлением структурирования проблемы на основе системного анализа и создания автоматизированных систем поддержки принятия решения, так как по мнению одного из основоположников искусственного интеллекта, академика Г.С. Поспелова автоматизация рутины –не прогресс, а регресс.

В условиях конкурентной (рыночной) экономики, в любой из отраслей – подотраслей – хозяйствующих субъектов на всех этапах жизненного цикла должна решаться общая задача выбора оптимального варианта всех соответствующих им организационных, технологических, технических, конструктивных составляющих материальных технологий, обеспечивающего получение максимального дохода при минимизации затрат. Т.е. в практике функционирования каждой отрасли стоит задача многокритериальной оптимизации, решение которой может осуществляться только на основе информационных технологий.

Наиболее актуальна информатизация в отраслях с большими удельными затратами бизнес-процессов получения производимой продукции, что делает проблему повышения их эффективности наиболее острой. В том числе, это относится к сельскохозяйственному растениеводству, имеющему удельные затраты 500-1000 \$/га, что связано с большой площадной распределенностью и высокой энергоемкостью обработки почвы, как в технологических операциях выращивания сельскохозяйственных растений, так и при создании систем обеспечения их факторов жизни. В частности, большие удельные затраты требуются при создании (1500-2000 \$/га и более) и поддержании в работоспособном состоянии мелиоративных систем, обеспечивающих один из важнейших урожаеобразующих факторов – формирование требуемого сельскохозяйственным растениям благоприятного водного режима непрерывно на протяжении всего вегетационного периода. В сочетании с необходимостью осуществления мелиоративных работ по реконструкции эксплуатации и управлению водным режимом на 2,9 млн. га осушенных земель, ежегодно требуются большие объемы капиталовложений и чрезвычайно актуальна проблема их эффективного использования. Несмотря на совершенствование технических средств механизации мелиоративных работ, удельные затраты на их осуществление неизбежно остаются большими, что вызывается энергоемкостью перемещения почвы, являющегося основной операцией создания конструктивных решений, соответствующих традиционным методам и способам отвода гравитационной воды осушительными системами (прокладка и подчистка открытой сети каналов, ложбин, воронок, планировка и профилирование поверхности, укладка дренажа).

Учитывая продолжительный период развития базовых отраслей, в частности сельскохозяйственной мелиорации, традиционные методы совершенствования их технологических (материальных) процессов, конструктивных решений, методов, способов на макроуровне практически задействованы. Качественное улучшение материальных технологий, конструктивных решений, более эффективных по соотношению «цена–качество»), возможно при переходе на микроуровень (био-, нано технологические решения и т. п.), что в обозримой перспективе представляется маловероятным.

В этих условиях информационные технологии являются важнейшим средством повышения эффективности существующих материальных технологий, посредством создания систем поддержки принятия решений, обеспечивающих выбор оптимальных вариантов структуры, параметров, режимов из имеющегося множества альтернатив при планировании, проектировании, управлении использованием всех урожаеобразующих факторов (водного, пищевого, химического, теплового режимов).

Фактически в производящих продукцию отраслях, в том числе сельскохозяйственном растениеводстве и важнейшей его составляющей – мелиорации, цифровая экономика означает качественно новый уровень оптимизации при планировании и проектировании бизнес-процессов на всех этапах жизненного цикла: строительства – эксплуатации – реконструкции, без которого невозможен переход от вала к прибыли, что в современной конкурентной экономике является безальтернативным условием повышения эффективности и развития от отдельного хозяйства до отрасли в целом.

В мелиорации система поддержки принятия решений базируется на оптимизационно-имитационных экономико-математических моделях, с более строгим описанием физических процессов функционирования мелиоративных систем во взаимосвязи неразрывного процесса динамики воды в цепочке мелиоративная сеть, мелиорированная почва – растения, связанных с ней других урожаеобразующих факторов формирования урожайности, на качественно новом уровне математического аппарата, соответствующего вычислительным возможностям современных компьютерных средств.

Другой составляющей системы принятия решений является подсистема информационного обеспечения оптимизационно-имитационной модели на основе мониторинга параметров и переменных состояния каждого конкретного объекта и его элементов, в том числе с использованием ДДЗ, с автоматизированным хранением данных в ГИС и БД. При этом наибольшую актуальность, как и в других технических системах [2], имеет автоматизация проектирования.

Результаты исследований. Первоначальное строительство мелиоративных систем в 50-80 гг., в связи с имевшим место дефицитом продовольствия, осуществлялось на цели хозяйственно-технического уровня общности (максимум урожайности и, соответственно, максимум быстрогодействия мелиоративных систем на осушение и увлажнение). При этом, в условиях домашней математики проектирование вынужденно осуществлялось с использованием предельно упрощенных зависимостей осушительного действия мелиоративной сети, основываясь на нормативно-типизированном одновариантном расчете параметров систем на усредненный по площади объекта нормативно принимаемый модуль стока расчетной обеспеченности критического периода [3], что вело к большим коэффициентам запаса параметров сети и, соответственно, ее капиталоемкости.

В условиях конкурентной экономики целью систем производства продукции любого вида, в том числе сельскохозяйственной, является экономическая эффективность. В связи с этим, в современных условиях цели (критерии) сельскохозяйственного производства, и соответственно одной из важнейших его составляющих – мелиорации, кардинально изменяются, в сравнении с условиями дефицита продовольствия, имевшими место в период первоначальной мелиорации.

В соответствии с прежней целью максимизации производства сельскохозяйственной продукции, первоначальная широкомасштабная мелиорация осуществлялась с ориентацией на обеспечение водного режима, максимально соответствующего требованиям сельскохозяйственных растений на протяжении периода вегетации. В соответствии с этим основной концепцией было создание все более технически «совершенных», – имеющих максимальное быстрогодействие на осушение-увлажнение мелиоративных систем, посредством проектирования параметров их элементов (каналов, дренажа) с повышенными коэффициентами запаса, использования

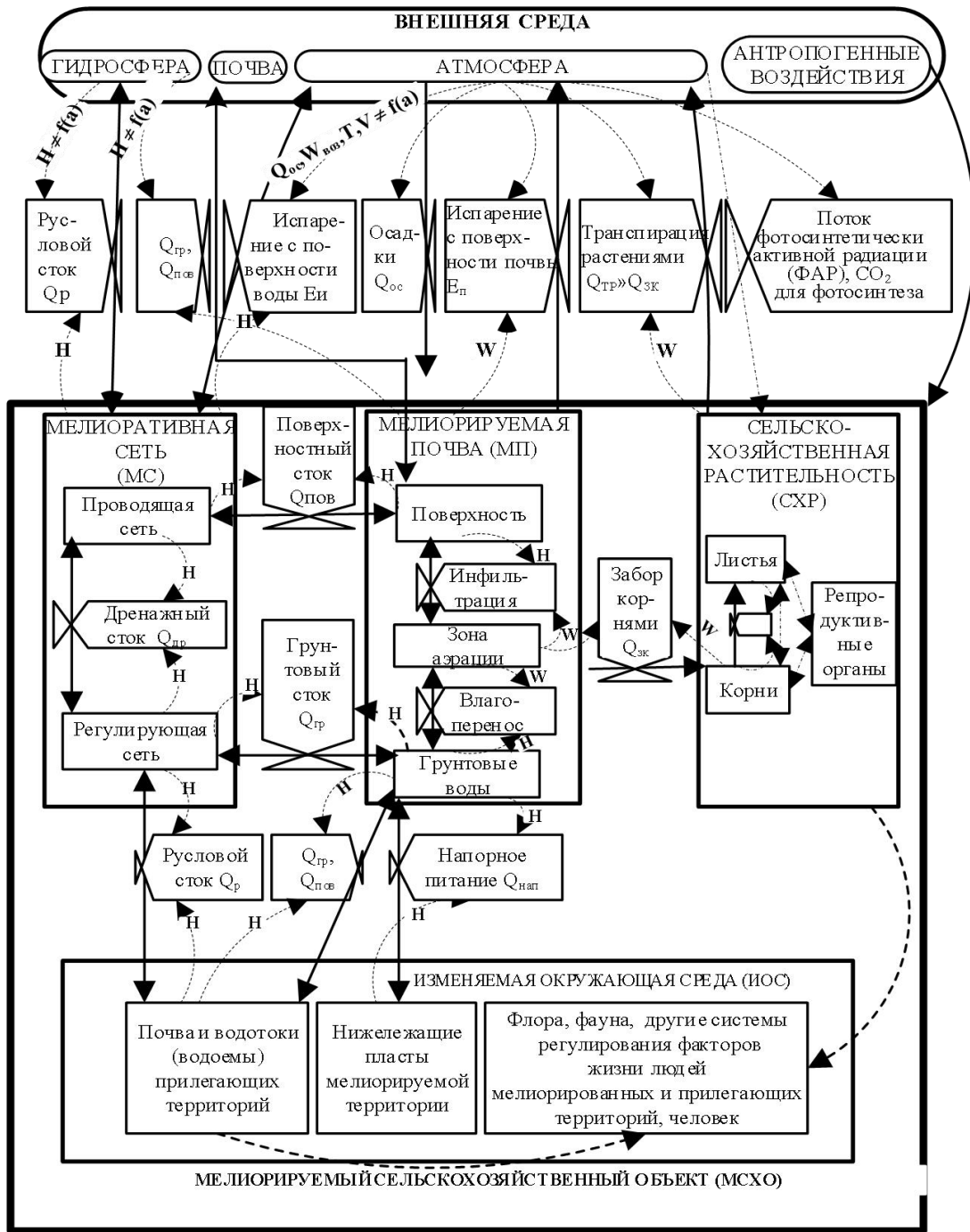
машинного водоподъема, реализации автоматизированного управления водным режимом и т.п. Очевидно, что повышение точности регулирования водного режима за счет все большего технического совершенствования систем ведет к росту удельных затрат. При этом рост урожайности от обеспечения более оптимального водного режима с каждым шагом повышения точности регулирования все более замедляется (становится все медленнее окупаемым).

В современных условиях конкурентной экономики, требующей от сельского хозяйства перехода с валовых показателей на экономические, главной задачей в мелиорации является изменение основополагающего подхода в принятии решений при планировании и проектировании мелиоративных мероприятий (строительство, реконструкция, ремонт мелиоративных систем, управление водным режимом) с целей технической работоспособности систем, к выбору вариантов мелиоративных объектов, отвечающих целям экономико-экологического уровня общности: максимум доходов, минимум затрат, минимум негативного влияния на окружающую среду [4]. В конкретных задачах эти цели могут выступать в виде критериев оптимальности или ограничений.

Анализ достижения сформулированного набора целей не может быть реализован при рассмотрении только непосредственно мелиоративной системы, определяющей исключительно величину затрат, а по неизбежности требует моделирования ее в составе более общего иерархического уровня общности.

Исходя из системного подхода, в качестве функционально полной системы должен рассматриваться мелиоративный сельскохозяйственный объект (МСХО), включающий четыре подсистемы: мелиоративная сеть (МС), мелиорируемая почва (МП), сельскохозяйственная растительность (СХР), изменяемая окружающая среда (Рис.1), неразрывно связанные единой системой целей экономико-экологического иерархического уровня общности и процессами взаимодействия, основополагающим среди которых является процесс динамики воды. На рис 1 динамика материального потока воды в подсистемах и водообмен между ними отображается с использованием диаграммной символики Форрестера, отражающей взаимосвязь переменных типа потока (расход воды) и типа уровня(напор), в русловых потоках сети каналов и дренажа подсистемы мелиоративная сеть, грунтового и поверхностного стока в подсистеме мелиорируемая почва и водопотребления в подсистеме сельскохозяйственная растительность с учетом сложной транзитивной цепочки прямых и обратных связей.

Схема Рис.1 отражает модель структуры и физическое представление технологического процесса агропроизводства на мелиорированных землях, формализация которого представляет матфизическую модель МСХО.



- - системы (подсистемы)-емкости, характеризующиеся уровнем;
- ▭ - функции решений; ○ - внешняя среда (надсистема "природа");
- - вещественные потоки (воды а также переносимых ею тепла, питательных элементов, химических элементов и т.д.);
- ⋯→ - потоки ФАР, CO₂, ассимилятов растений; - - - - - → - потоки информации;

$H(W)$ - переменные типа уровня (уровень воды(влажность) в соответствующей подсистеме; Q - переменные типа потока; $W_{воз}, T, V$ - влажность, температура воздуха, скорость ветра; a - альтернативы (варианты структуры, изменяемых параметров, управлений) в МСХО.

Рис.1 Модель взаимосвязи процессов динамики воды на мелиорируемом сельскохозяйственном объекте и прилегающей территории

Средняя площадь отдельного МСХО составляет около 600 гектар, на равнинных землях Полесской низменности достигая 1000 гектар и более. Общее число отдельных

МСХО составляет 4950 штук, с генерализацией до мелиоративного комплекса Беларуси в целом, представленных на рис.2.



Рисунок 2. Мелиоративный комплекс Республики Беларусь

Всего в Беларуси в результате проведенной широкомасштабной мелиорации было осушено и находится в сельскохозяйственном использовании около 3 млн. га земель – треть всех сельхозугодий. Мелиоративные системы были построены на протяжении длительного периода (основная часть в 50-90 годы), в связи с чем в настоящее время находятся в различном состоянии. В результате этого на них имеет место необходимость в мероприятиях, соответствующих различным этапам жизненного цикла, реализация которых осуществляется в рамках текущей государственной программы «Аграрный бизнес» на 2021-2025 годы (рис.3).

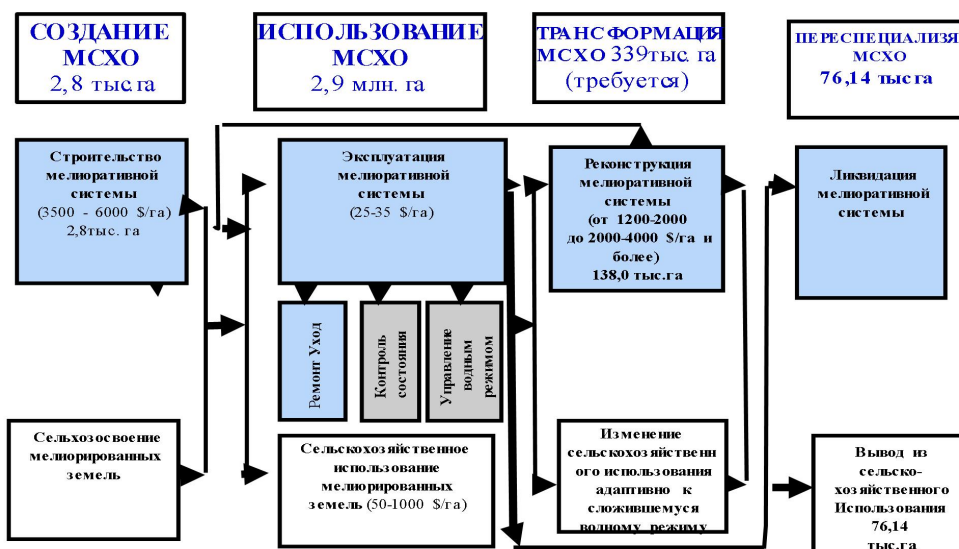


Рисунок 3. Жизненный цикл мелиоративных сельскохозяйственных объектов (МСХО) отрасли и его этапы в Государственной программе «Аграрный бизнес» на 2021-2025 гг.

Переход с подходов, основанных на критериях работоспособности систем к экономической эффективности в особенности актуален для наиболее затратных этапов жизненного цикла мелиоративных систем: нового строительства реконструкции, капитальных ремонтов.

Решение этой задачи может быть обеспечено только на основе разработки и использования в мелиорации и сельхозиспользовании мелиорированных земель инновационных решений как в материальных, так и в играющих основную роль при принятии решений, информационных технологиях.

Основополагающими концептуальными подходами для достижения этого являются:

– системный подход: планирование мелиоративных мероприятий от общего к частному: республиканский уровень – бассейновые схемы – конкретные мелиоративные объекты. Рассмотрение при проектировании в качестве единой оптимизируемой системы мелиоративного сельскохозяйственного объекта (МСХО);

– использование адаптивных стратегий: рассмотрение при выборе решений в качестве возможных альтернатив наряду с исключительно активными природо-изменяющими стратегиями («кардинальной» мелиорации), также и менее затратных и оказывающих меньшее воздействие на окружающую среду, вариантов приспособления сельскохозяйственного использования земель, проходимости используемой сельскохозяйственной техники, выбираемой интенсивности осушения наиболее пониженных участков МСХО, к складывающемуся водному режиму. Особо следует отметить, что при этом понятие адаптации следует понимать не относительно хозяйственно-функциональных характеристик, а по экономическим показателям:

– переход от нормативно типизированных решений к многовариантным расчетам на основе компьютерных систем поддержки принятия решения базирующихся на оптимизационно-имитационных экономико-математических моделях с распределенными параметрами, адекватно описывающих физику процессов осушительного действия, с использованием объективных данных о реальном состоянии систем, для обеспечения «Точной мелиорации» – осуществления исключительно реально обоснованных мелиоративных мероприятий с действительно необходимыми параметрами дифференцированными по площади объекта (т.е. исключения непродуктивных затрат);

– переход в проектировании, при описании основополагающего мелиоративного процесса – осушительного действия мелиоративной системы, от использования предельно упрощенных, ориентированных на возможность осуществления ручного счета, зависимостей течения воды в грунте и мелиоративной сети (установившееся, стационарное, равномерное движение воды с эвристически принимаемым расходом расчетной обеспеченности расчетного периода), к многовариантному автоматизированному проектированию, базирующемуся на физически адекватных реальным условиям осушения нестационарным моделям с распределенными параметрами;

– оптимизация вида и интенсивности сельхозиспользования мелиорированных земель на основе надежных многофакторных мультипликативных моделей урожайности сельскохозяйственных культур;

– организация эффективной системы автоматизированного информационного обеспечения принятия решений с распределенными параметрами при планировании и проектировании мелиорации и сельхозиспользования мелиорированных земель на всех иерархических уровнях их осуществления, на основе проведения агро-гидро-эколого-мелиоративного мониторинга параметров и переменных состояния, ведения книги истории полей, контроля и диагностики всех подсистем конкретного мелиоративного сельскохозяйственного объекта и его элементов, в том числе с использованием данных дистанционного зондирования земли, средств неразрушающего контроля с автоматизированным хранением, обработкой и использованием результатов наблюдений в

реализованной на основе геореляционной концепции пространственно-временной информационной системе, базирующейся на интегрированном использовании ГИС и БД;

- максимальное использование работоспособных элементов при реконструкции мелиоративных систем;
- разработка конструкций мелиоративных систем и их элементов с использованием новых материалов;
- разработка многофункциональных энергосберегающих технических средств, сельскохозяйственных машин повышенной проходимости.
- эффективная эксплуатация МСХО, обеспечивающая максимальный промежуток лет между дорогостоящими реконструкциями.

Принятие эффективных решений на всех уровнях планирования, проектирования, строительства, эксплуатации мелиоративных объектов и их сельхозиспользования требует использования систем поддержки принятия решений, разработанная структура которых, включающая оптимизационные модели и их информационное обеспечение приведена на рис.3.

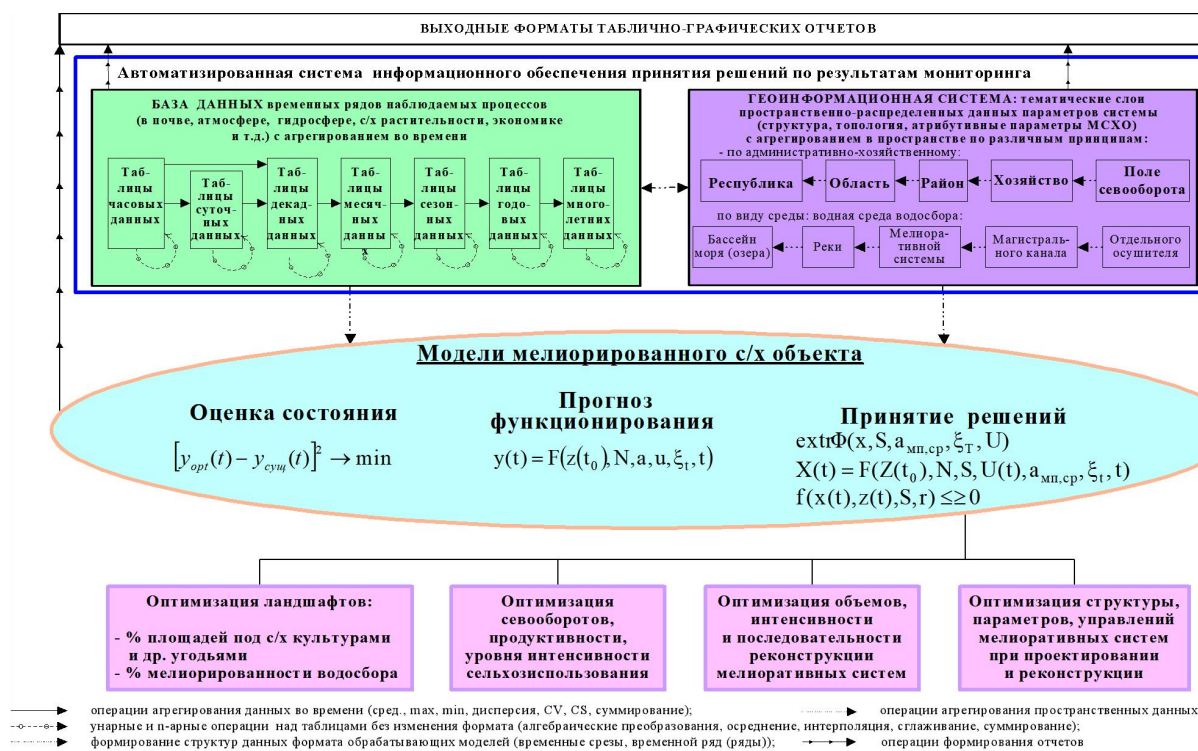


Рисунок 4. Система поддержки принятия экономико-экологически обоснованных решений в мелиорации и сельхозиспользовании мелиорированных земель с геореляционной структурой информационного обеспечения по результатам мониторинга агро-гидро-гео-метео-показателей объектов

На основании проведенного системного анализа для современных условий конкурентной экономики, сформулирована необходимость оптимизации мелиоративных мероприятий, на цели экономико-экологического уровня общности (максимум доходов, минимум затрат, экологические ограничения на параметры и процессы), прежде всего, при проектировании наиболее дорогостоящих реконструкции и строительства новых систем [4]. Учитывая взаимосвязь степени регулирования водного режима (затраты) с урожайностью сельскохозяйственных культур (доходы), прослеживание целей экономического уровня требует рассмотрения в качестве единой оптимизируемой

системы мелиоративного сельскохозяйственного объекта (МСХО), включающего подсистемы: мелиоративная сеть, мелиорируемая почва, сельскохозяйственная растительность, изменяемая окружающая среда, неразрывно взаимосвязанные транзитивными цепочками прямых и обратных связей протекающих в них процессов взаимодействия, формирующихся под воздействием стохастических погодно-климатических воздействий.

Разработанная методика оптимизации параметров по критериям экономико-экологического уровня (рисунок 2) заключается в выборе наилучшего варианта на основе многовариантных расчетов на модели МСХО. При этом, наряду с активными природоизменяющими стратегиями регулирования водного режима под требования растений (варианты структуры и параметров мелиоративных систем), в качестве альтернатив рассматриваются и адаптивные стратегии – смена вида и интенсивности сельхозиспользования под естественно-природные условия и водный режим, сложившийся в результате износа системы в процессе многолетнего использования, что обеспечивает выбор согласованных по экономическим критериям параметров мелиоративных систем и сельхозиспользования [5]. Наиболее актуально использование адаптивных подходов (севообороты с влаголюбивыми культурами и сортами, сельскохозяйственная техника повышенной проходимости) в течение продолжительного периода на площадях, по результатам инвентаризации требующих реконструкции, но не вошедших, из-за ограниченности финансирования, в программу работ на текущую пятилетку (в действующей программе это половина площадей из 350 тыс. га требующих реконструкции).

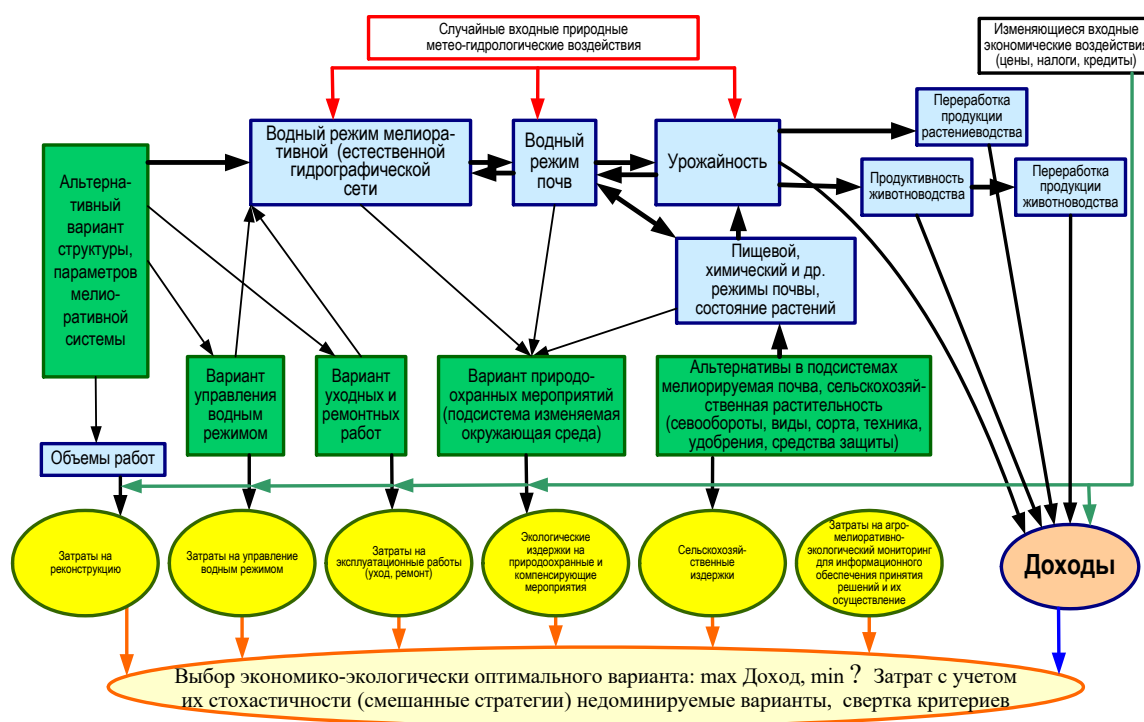


Рисунок 5 – Методология проектирования реконструкции мелиоративных сельскохозяйственных объектов на цели экономико-экологического уровня общности

Исходя из необходимости учета затрат всех этапов жизненного цикла, сформулировано требование о целеориентированности выбора оптимального варианта самой системы принятия решений из возможных альтернатив (от наиболее строгих уравнений с распределенными параметрами со случайными стохастическими входными

погодно-климатическими воздействиями, до наиболее упрощенных моделей типа «черный ящик» и производственные функции), основанного на согласованном выборе уровня точности модели принятия решения и необходимого для расчета на ней информационного обеспечения, при котором имеет место наибольшее превышение дохода от прибавки урожайности в результате более точного принятия решения (выбора оптимальной структуры и параметров при проектировании мелиоративной системы, алгоритма управления при эксплуатации) над затратами на сбор информации и реализацию расчетов.

Аналогично этому сформулированы альтернативные варианты методов контроля состояния мелиоративных систем по: конечным экономическим характеристикам (прибыль, рентабельность) – хозяйственным (урожайность) – переменным состояниям (водный режим почвы – водный режим мелиоративной сети) – параметрам мелиоративной сети, имеющих различный уровень точности и затрат на осуществление.

Для реализации вышеприведенной системы принятия решений разработан комплекс взаимосвязанных программных средств, предназначенных для оценки состояния, прогноза функционирования систем и их оптимизации на основе многовариантных расчетов.

Информационным ядром системы является реализованная на основе геореляционной методологии автоматизированная система, включающая разработанную специализированную базу данных временных рядов агро-метео-эколого-гидромелиоративного мониторинга, обеспечивающую хранение имеющих быстрое характерное время изменения (часы) переменных состояния водного режима и других урожаеобразующих факторов во всех подсистемах МСХО (уровни, расходы, влажность и др.) и входных случайных агрометеорологических факторов (осадки, температура и т.д.), интегрированную с ГИС, обеспечивающей хранение структуры и параметров МСХО. Так как система предназначена не просто для хранения данных (наджность хранения обеспечивается по умолчанию), а для эффективного автоматизированного информационного обеспечения разнообразных задач прогноза, оценки, оптимизации, то разработан ряд программных средств по обработке, пространственно-временному анализу собранных в автоматизированной системе данных. Функционал системы включает синхронизацию, расчет различного вида статистик агрометеофакторов любого уровня агрегирования, их аппроксимацию во времени и зонирование по территории, алгоритм расчета скользящего среднего для прослеживания направленности трендов изменения агрометеофакторов и урожайности в последовательном ряду лет, расчет обеспеченности урожайности сельхозкультур, изменяющейся в связи с колебанием по годам погодно-климатических условий, с сохранением результатов в виде наборов данных различной структуры для использования в различных приложениях [5-7].

Используя собранные данные мониторинга агрометеофакторов за весь имеющийся период наблюдений Гидрометслужбы и урожайности различных сельскохозяйственных культур в сортоиспытаниях на сортоиспытательных станциях Государственной инспекции по испытанию и охране сортов растений, с использованием геоинформационных технологий, осуществлен пространственно-временной анализ и получены выводы о том, что в годы с различной обеспеченностью урожайности наибольшую ее величину имеют различные культуры, т.е. культуры с наибольшей урожайностью в среднемноголетние годы и имеющие наибольшую ее величину в годы с экстремальными агрометеоусловиями (осадки, температура), могут различаться.

В соответствии с рис.5 разработан и реализован комплекс программных средств, предназначенных для решения задач контроля оценки, прогноза и оптимизации мелиоративных систем и сельскохозяйственного использования земель, а также организации их автоматизированного информационного обеспечения, взаимосвязь которых приведена на рис.6.



Рисунок 6. Комплекс программных средств системы поддержки принятия решений в мелиорации

Исходя из того, что эффективность функционирования мелиоративных объектов определяется рядом характеристик (водный режим мелиоративной сети, мелиорированных земель, метеофакторы, плодородие почвы и т.п.), пространственно-распределенных на больших площадях и изменяющихся во времени (в течение периода вегетации и по годам) разработана геореляционная структура хранения результатов их мониторинга. Она включает хранение пространственно распределенной структуры и параметров мелиоративных объектов в тематических слоях геоинформационной системы, а переменных быстроизменяющихся процессов (температура, осадки, уровни, расходы и т.д.) – в разработанной специализированной реляционной базе данных временных рядов «GEO2», с организацией интерфейса между ними].

В настоящее время в БД собрано более 15 млн. данных по метеостанциям Беларуси, и других объектов, являющихся надежным информационным обеспечением различных практических задач мелиорации. С помощью реализованных клиентских приложений обеспечиваются разнообразные обработки временных рядов, их анализ, графическое и табличное представление.

Единая методология информационного обеспечения всех рассматриваемых задач (программно реализованных моделей) обеспечивает возможность беспрепятственного автоматизированного обмена данными между информационными подсистемами, необходимыми различным субъектам мелиоративной деятельности: подсистема госучета мелиоративных объектов в генерализованных показателях «Аграрный бизнес» для разработки программ мелиорации республиканскими и областными органами управления; подсистема обеспечения бизнес-процессов в мелиоративных предприятиях (паспортизация мелиоративных систем) «Аграрный бизнес» планирования работ по эксплуатации, реконструкции, контроль их выполнения и технического состояния мелиоративных систем по ежегодным обследованиям и инвентаризациям; книга истории полей агропредприятия и практически являющаяся ее составной частью, система мониторинга почв по турам почвенных и агрохимических обследований, обеспечивающая отдельный учет урожайности на мелиорированных землях для планирования их сельскохозяйственного использования, обоснованного выбора объектов требующих реконструкции;

система мониторинга агро-гидрометеорологических данных для информационного обеспечения проектирования. В результате этого все вышеперечисленные подсистемы фактически являются частями единой информационной системы агропроизводства. При этом исключается непродуктивный повторный сбор и ввод данных.

Информационное ядро системы (ГИС, БД) обеспечивает автоматизированное формирование любых пространственно-временных наборов данных различной структуры, необходимых для оптимизационно-имитационных моделей с распределенными параметрами осушительного действия мелиоративных систем, формирования поверхностного стока (*VRLAB*, ГИС модели поверхностного стока, книга истории полей севооборота (рис.4)), а также других наборов данных, необходимых для любых задач (используемых моделей) на различных иерархических уровнях управления.

Заключение. Разработанная система поддержки принятия решений обеспечивает повышение эффективности мелиорации за счет осуществления многовариантных расчетов для выбора оптимальных вариантов при планировании и проектировании реконструкции мелиоративных систем и сельскохозяйственного использования мелиорированных земель, основываясь на реальных данных.

Список литературы

- [1] Государственная программа «Цифровое развитие Беларуси» на 2021-2025 годы. Постановление СМ РБ от 02.02.2021 № 66
- [2] Моисеев Н.Н. Математические задачи системного анализа –Москва: Наука,1981. – 487 С.
- [3] Технические условия и нормы проектирования осушительных систем в Белорусской ССР/Белгипроводхоз-Минск, 1970. -330с.
- [4] 4.Вахонин, Н. К. Мелиоративное проектирование с позиций системного анализа / Н. К. Вахонин // Прогнозы водного режима при мелиорации земель. –Минск: БелНИИМиВХ, 1988. – С. 47-61.

MODELING AND FORMATION OF SPATIO-TEMPORAL INFORMATION SUPPORT WHEN INFORMATIZING, USING THE EXAMPLE OF THE HYDRORECLAMATION INDUSTRY

Vakhonin N.K.

*Professor of the department
computer information
technologies MITSO, candidate
technical sciences, associate professor*

Annotation. The article presents the results of a system analysis of the problem of developing a decision support system in large, spatially distributed natural-technical systems, with stochastic input influences, using the example of informatization of the reclamation industry. A complete hydroreclamation system is identified, connected by common economic and environmental goals and the inseparable process of water dynamics an agricultural facility subject to optimization at all stages of its life cycle (design, creation, use, control, reconstruction). Alternative options for models of system functioning and their information support are formulated. It is shown that the choice of their optimum is itself a task of goal-oriented decision making, which predetermines the assessment of the effectiveness of informatization based on the results of its use in the real sector of the economy. A number of information tools developed for land reclamation are presented.

Key words: system analysis, subject area, real sector of the economy, large systems, material systems, hydroreclamation industry, hydroreclamation agricultural facility, informatization, digital economy, decision support system, optimization, goals, criteria, economic efficiency, stages of the systems life cycle, design of information systems, diagrammatic symbols, model of system functioning, equations with distributed parameters, black box model, neural network model, stochastic input influences, spatiotemporal information support, geo-relational databases, big data, data rotation, geographic information systems.

УДК 004.021:004.75

АВТОНОМНАЯ ПОСАДКА БПЛА С ИСПОЛЬЗОВАНИЕМ ВИЗУАЛЬНОГО СЕРВОУПРАВЛЕНИЯ



В.Ф. Алексеев

Доцент кафедры проектирования информационно-компьютерных систем БГУИР, кандидат технических наук, доцент
alexvikt.minsk@gmail.com



Е.И. Бавбель

Аспирант кафедры проектирования информационно-компьютерных систем БГУИР
egorigorevichw9@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.

Е.И. Бавбель

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации.

Аннотация. Автономная посадка беспилотных летательных аппаратов (БПЛА) играет важную роль в продвижении беспилотных систем к практическому применению. Однако традиционный метод управления посадкой на основе местоположения в значительной степени опирается на спутниковые системы навигации и позиционирования, которые нельзя применять в условиях слабых спутниковых сигналов, таких как помещения и долины. Для решения этих проблем в данной статье предлагается метод визуального сервоуправления автономной посадкой БПЛА на основе программного обеспечения Aerostack. Этот метод принимает изображение бортовой камеры в реальном времени в качестве входных данных и извлекает особенности меток на земле. Кроме того, создается закон визуального сервопривода автономной посадки для получения сигналов управления скоростью, которые могут точно направить дрон на посадку. По результатам экспериментов проводимыми за последние годы, показывают, что предложенный алгоритм может обеспечить надежную и точную автономную посадку без сигналов спутниковой навигации.

Ключевые слова: Автономная посадка, сервопривод, БПЛА.

Введение. В последние годы беспилотные летательные аппараты получили широкое распространение как в военной, так и в гражданской сферах из-за их потенциальной ценности в практическом применении. Автономная посадка БПЛА является одним из жизненно важных методов продвижения развития БПЛА. Многие приложения, такие как спасение при стихийных бедствиях, транспортировка грузов, требуют точной автономной посадки БПЛА для выполнения миссий [1]. Традиционный метод управления посадкой на основе местоположения в значительной степени опирается на спутниковые системы навигации и позиционирования. Однако существующие системы спутниковой навигации и позиционирования, такие как *GPS* и *BeiDou*, недостаточно точны в гражданской сфере, и навигационный сигнал легко теряется при наличии

препятствий над БПЛА, не говоря уже о закрытых помещениях и долинах. В связи с этим для решения этой проблемы очень эффективен метод визуального сервоуправления [2], который генерирует управляющий сигнал на основе визуальной обратной связи. Кроме того, поскольку датчики, используемые при визуальном сервоконтроле, в основном являются визуальными, измерение окружающей среды является бесконтактным. В то же время по сравнению с традиционным методом навигационного управления можно получить больший объем информации об окружающей среде и эффективно повысить точность управления. Исследования показали, что целым рядом авторов предлагается различные подходы к решению проблемы автономной посадки БПЛА. Тем не менее, эти подходы не учитывают масштабируемость интеграции других задач, таких как отслеживание маршрута и отслеживание целей. Для обеспечения высокого уровня автономности желательна масштабируемость поддержки нескольких миссий.

Авторами рассматриваются программное обеспечение *Aerostack* – архитектурная основа беспилотной авиационной системы. Эта архитектура основана на концепции модульности, позволяющей легко расширять другие миссии и другие виды платформ. Таким образом, на основе структуры *Aerostack* в данной статье предлагается подход к визуальному сервоуправлению автономной посадкой БПЛА. Такой подход не только обеспечивает полностью автономную и точную посадку, но также использует масштабируемость *Aerostack*. По результатам моделирования в данной среде можно сказать, что подход позволяет обеспечить надежную и точную автономную посадку.

Анализ принципов работы сервоуправления БПЛА. Полетная миссия дрона обычно включает в себя три этапа: взлет, выполнение конкретной задачи и последующая посадка. Общий процесс визуального сервоуправления автономной посадкой БПЛА показан на рисунке 1. Чтобы проверить эффективность автономной посадки авторами показан пример управления дроном таким образом, чтобы он направлялся к целевой области перед приземлением, и гарантировал, что координаты положения дрона будут подчиняться равномерному случайному распределению по площади.

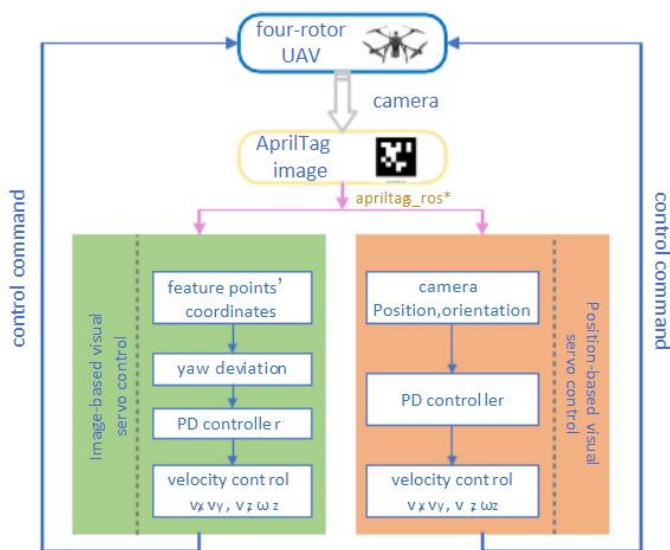


Рисунок 1. Процесс визуального сервоуправления

На протяжении всего проекта направленная вниз камера является основным источником данных для автономной посадки. Существует два способа реализации визуального сервопривода: визуальный сервопривод на основе положения (*PBVS*) и визуальный сервопривод на основе изображения (*IBVS*). Следует отметить, что алгоритмы управления в этих двух режимах не ограничены. Идеальная цель успешного приземления состоит в том, чтобы опорная планка касалась земли, а направленная вниз камера была направлена на метку на земле. Если камера

установлена непосредственно под дроном, опорная планка заблокирует дрон и закроет поле зрения камеры. В этом случае метка не будет распознана, поэтому рекомендуется установить камеру под правой рукой дрона. Когда дрон успешно приземляется, он не находится непосредственно над меткой, что мало влияет на оценку алгоритма миссии и управления.

Основываясь на платформе *Aerostack*, возможно расширить функциональность модуля визуального сервоуправления и реализации моделирования в *Gazebo* [3]. На рисунке 2 представлен проект архитектуры реализации автономной посадки БПЛА на основе визуального сервоуправления. В процессе автономной посадки камера, направленная вниз, захватывает изображение, а средство извлечения признаков извлекает особенности метки на земле. Контроллер движений генерирует сигналы управления скоростью для движения дрона на основе информации о функциях, тем самым управляя дроном так, чтобы он точно приземлился в заданное наземное положение.

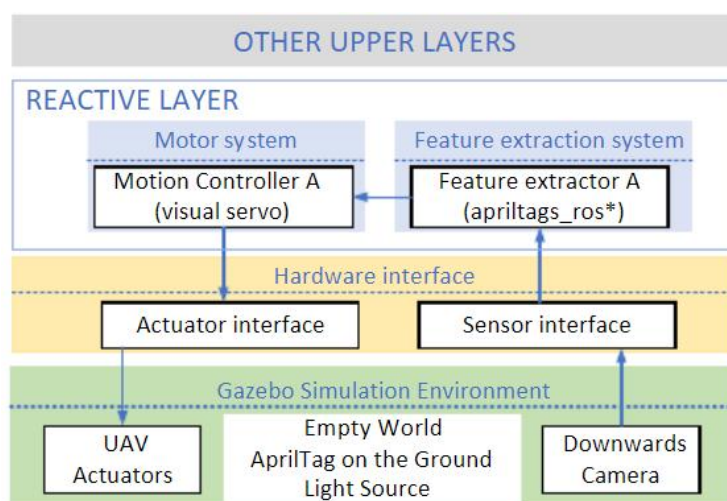


Рисунок 2. Проектирование архитектуры в *Aerostack*

Модуль визуального сервоуправления принадлежит к *Reactive Layer* структуры *Aerostack*. Экстрактор функций получен на основе пакета с открытым исходным кодом. Чтобы получить управляющие переменные, необходимые для визуального сервоуправления на основе изображения, необходимо добавить функцию публикации пиксельных координат центральной точки и четырех угловых точек тега (в данной статье пакет называется *apriltags_ros**). Стоит отметить, что в модулях расширенных функций используется метод связи узла *ROS* – *Topic*. Контроллер в системе может использовать различные функции, предоставляемые экстрактором функций, для генерации сигналов управления движением дрона с использованием другого алгоритма управления.

Создание среды моделирования. В *Gazebo* модели (рисунок 3), которые необходимо создать, – это в основном пустой мир (*AprilTag* на земле, источники света, БПЛА и камера, направленная вниз). Источники света, используемые в данной статье, являются точечными и располагаются непосредственно над меткой.

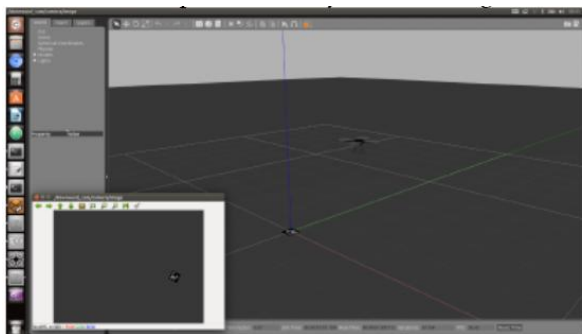


Рисунок 3. Эксперимент в Gazebo

Основное содержание симуляционного эксперимента заключается в загрузке проекта миссии верхнего уровня, написанного на языке *Python*, симуляции в среде *Gazebo* и получении экспериментальных данных. При автономной посадке БПЛА с использованием визуального сервоуправления, основанной на процессе автономной посадки *Aerostack*, вертикальная скорость БПЛА устанавливается равной 0,2 м/с, поэтому время, необходимое для фазы приземления, составляет примерно 20 секунд. Параметры *PD*-контроллера показаны на рисунке 4. Их можно настроить в соответствии с реальной ситуацией.

Table 1: PD Controller Parameters

PBVS	K_p	K_d	IBVS	K_p	K_d
x	0.9	0.005	u	0.001	0.00001
y	0.9	0.001	v	0.001	0.00001
yaw	0.3	0.001	yaw	0.01	0.0001

Рисунок 4. Параметры *PD*-контроллера

Визуальное управление сервоприводами на основе позиции. Приведены и проанализированы десять экспериментов по моделированию с использованием *PBVS* в различных условиях освещенности, отклонение положения приземления показано на рисунке 5 [4].

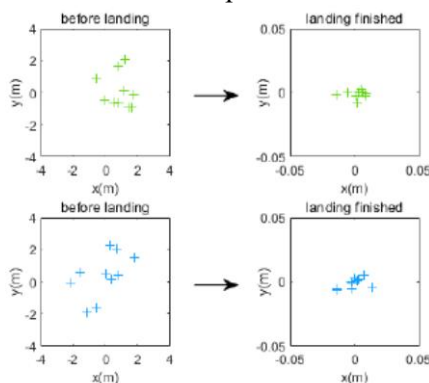


Рисунок 5. Отклонение позиции приземления (*PBVS*). Символы зеленого креста обозначают результаты экспериментов при 1-кратной интенсивности света, а синие – при 4-кратной интенсивности света

По теореме Пифагора можно подсчитать, что при 1-кратной силе света среднее отклонение положения при завершении приземления составляет 0,0069 м, а дисперсия – $1,0924 \cdot 10^{-5}$ м². Менее чем в 4 раза среднее отклонение положения составляет 0,0070 м, а отклонение – $3,0768 \cdot 10^{-4}$.

Угол отклонения от курса в начале и конце приземления показан на рисунке 6. При 1-кратной интенсивности света среднее отклонение от курса при посадке составляет 0,0147 рад, а дисперсия составляет $3,1363 \cdot 10^{-4}$. Напротив, среднее приземление при 4-кратном отклонении света составляет 0,0111 рад, а дисперсия – $1,5730 \cdot 10^{-4}$.

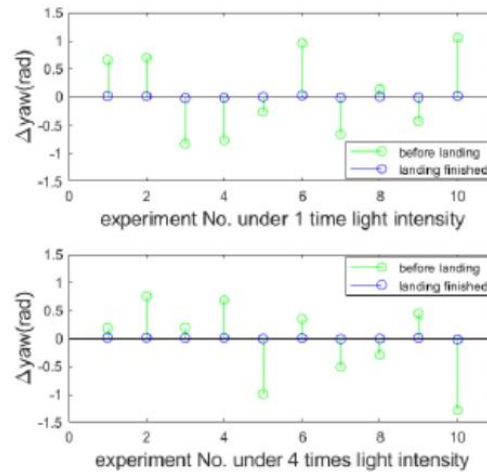


Рисунок 6. Отклонение от курса при посадке

Чтобы более интуитивно объяснить процесс приземления, отклонения положения и угла рыскания показаны на рисунке 7. Здесь показаны три приземления.

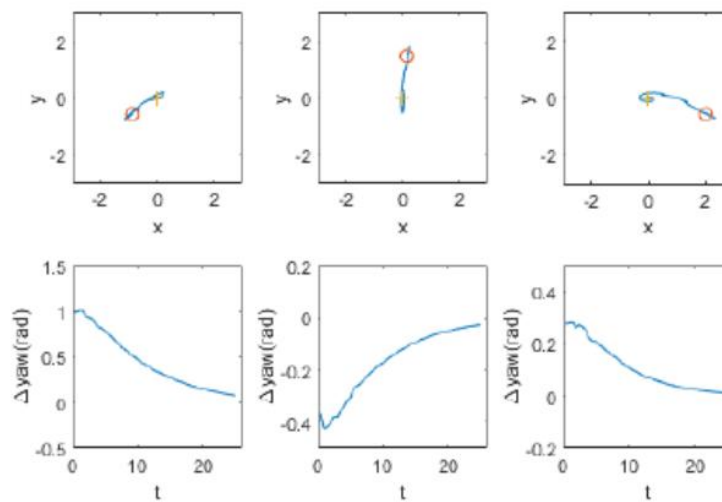


Рисунок 7. Отклонения положения и угла отклонения от курса в процессе приземления во временных рядах (PBVS). Круги относятся к начальным точкам, а символы «+» относятся к конечным точкам.

Визуальное управление сервоприводами на основе изображений. В тех же условиях моделирования, что и прежде приведем 10 экспериментов при различных условиях освещенности [4].

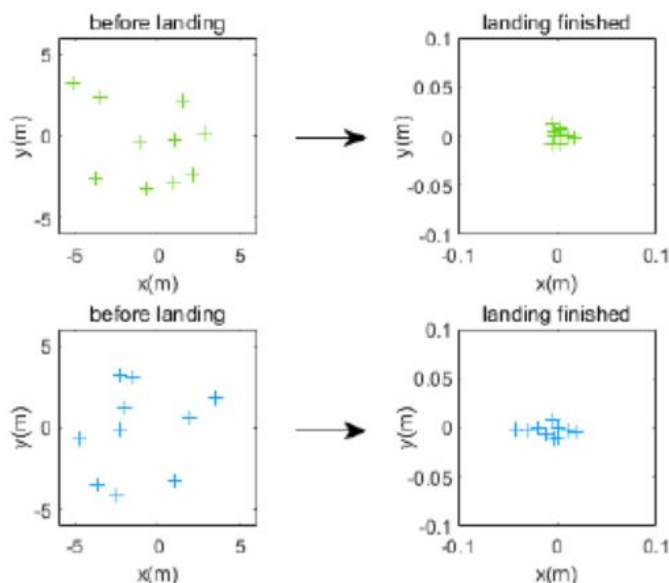


Рисунок 8. Отклонение посадочной позиции (IBVS). Зеленые крестики обозначают результаты экспериментов при 1-кратной интенсивности света, а синие – при 4-кратной интенсивности света

В соответствии с предыдущим данными мы можем получить отклонение положения, используя теорему Пифагора. Как показано на рисунке 8, при 1-кратной интенсивности света при приземлении отклонение положения составляет 0,0093 м, а отклонение составляет $2,2089 \cdot 10^{-5}$ м². При четырехкратной интенсивности освещения два результата составляют 0,0167 м и $1,4128 \cdot 10^{-4}$ м².

Отклонение угла рыскания в начале и конце приземления показано на рисунке 9. При 1-кратной интенсивности света среднее отклонение рыскания составляет 0,0044 рад, а дисперсия составляет $3,4383 \cdot 10^{-5}$. При четырехкратной интенсивности освещения среднее отклонение от курса составляет 0,0062 рад, а дисперсия – $4,3906 \cdot 10^{-5}$.

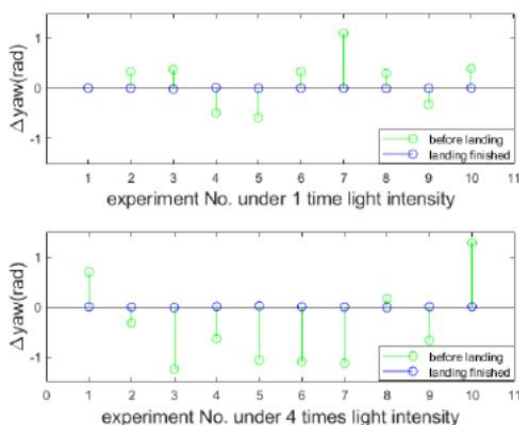


Рисунок 9. Отклонение от курса при посадке

Аналогичным образом, чтобы лучше описать процесс приземления, на рисунке 10 показаны изменения положения и угла отклонения от курса во время трех приземлений.

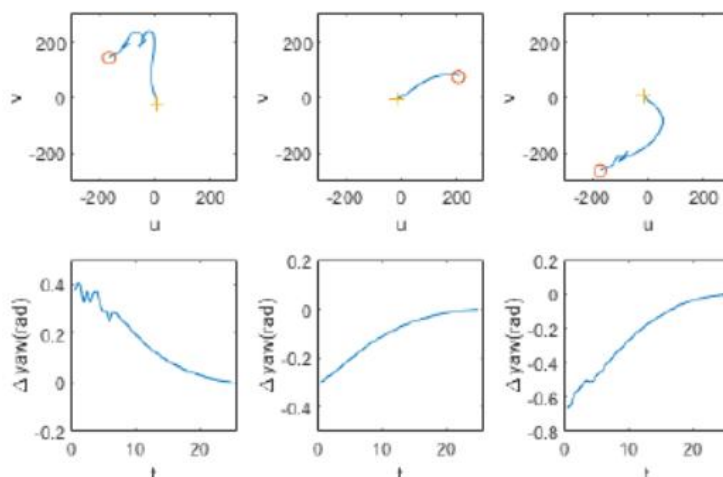


Рисунок 10. Отклонения положения и угла отклонения от курса в процессе приземления во временных рядах (IBVS). Кружки относятся к начальным точкам, а символы «+» относятся к конечным точкам

В симуляционном эксперименте алгоритмы визуального сервоуправления на основе положения и изображения показывают хорошую надежность и точность при автономной посадке дрона, что указывает на то, что автономная посадка достигается без спутниковой навигации и инерциальных навигационных датчиков. Но следует отметить, что в реальных приложениях все еще существует много проблем. В конце концов, среда моделирования не может полностью моделировать реальную среду. Например, в *Gazebo* невозможно смоделировать сложные световые и воздушные потоки, что может оказать большое влияние на устойчивость и надежность автономных площадок.

Заключение. Чтобы решить существующие проблемы неточного положения приземления и высокой зависимости от спутниковой навигации во время приземления БПЛА, в статье проведен анализ, и обработка данных полученные в ходе моделирования в программной среде и рассмотрен полностью автономный подход визуального сервоуправления приземлением БПЛА на основе платформы *Aerostack*. Основная работа на такой платформе включает в себя: проектирование миссии верхнего уровня, построение среды моделирования, расширение функций пакета *apriltags_ros*, добавление узла визуального сервоуправления на основе позиции и изображения в *Aerostack* и построение канала связи между *Aerostack* и *Gazebo*. Основываясь на этом, были проверена осуществимость общей конструкции. Предложенные алгоритмы визуального сервоуправления могут обеспечить надежную и точную цель приземления без спутниковой навигации.

Список литературы

- [1] Maza, I., Caballero, F., Capitán, J., Martínez-de-Dios, J. R., & Ollero, A. (2011). Experimental results in multi-UAV coordination for disaster management and civil security applications. *Journal of intelligent & robotic systems*, 61(1-4), pp. 563-585.
- [2] Chaumette, F., & Hutchinson, S. (2006). Visual servo control. I. Basic approaches. *IEEE Robotics & Automation Magazine*, 13(4), pp. 82-90.
- [3] Rodríguez-Ramos, A., Sampedro, C., Carrio, A., Bavle, H., Fernández, R. A., Milošević, Z., & Campoy, P. (2016, October). A monocular pose estimation strategy for uav autonomous navigation in gnss-denied environments. In *Proceedings of the International Micro Air Vehicle Conference and Flight Competition*, Beijing, China (pp. 17-22).
- [4] Guanzheng Wang, Zhihong Liu, and Xiangke Wang. 2019. UAV Autonomous Landing using Visual Servo Control based on Aerostack. In *Proceedings of the 3rd International Conference on Computer Science and*

Application Engineering (CSAE '19). Association for Computing Machinery, New York, NY, USA, Article 127, pp. 1–6.

Авторский вклад

Алексеев Виктор Федорович – руководство исследованием в области автономного полета, в частности визуального сервоуправления.

Бавбель Егор Игоревич – постановка задачи исследования метода визуального сервоуправления, анализ результатов моделирования на платформе *Aerostack* и последующей интеграции в *Gazebo*.

AUTONOMOUS UAV LANDING USING VISUAL SERVO CONTROL

V.F. Alexseev

E.I. Bavbel

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

*Postgraduate student, Department
of Information Computer Systems
Design BSUIR*

Abstract. Autonomous landing of unmanned aerial vehicles (UAVs) plays an important role in advancing unmanned systems to practical applications. However, the traditional location-based landing control method relies heavily on satellite navigation and positioning systems, which cannot be applied in environments with weak satellite signals, such as indoors and valleys. To solve these problems, this paper proposes a visual servo control method for autonomous UAV landing based on *Aerostack* software. This method takes the real-time airborne camera image as input and extracts the features of marks on the ground. In addition, a visual servo law of autonomous landing is created to obtain speed control signals that can accurately guide the drone to land. Based on the results of experiments conducted in recent years, they show that the proposed algorithm can provide reliable and accurate autonomous landing without satellite navigation signals.

Keywords: Autonomous landing, servo drive, UAV

УДК 004.021:004.75

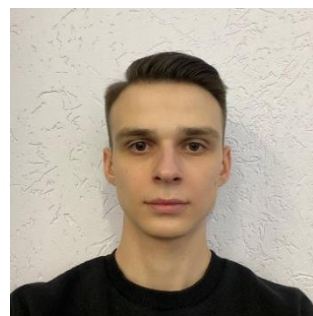
ОБРАБОТКА БОЛЬШИХ ДАННЫХ В ТУРИСТИЧЕСКОЙ ИНДУСТРИИ



С.А. Мигалевич
Магистр технических наук,
начальник центра
информатизации и
инновационных разработок
migalevich@bsuir.by



А.Н. Марков
Магистр технических
наук, заместитель
начальника центра
информатизации и
инновационных
разработок
a.n.markov@bsuir.by



А.Д. Стрельцов
Студент Белорусского
государственного
университета
информатики и
радиоэлектроники
a.streltsov@bsuir.by

С.А. Мигалевич

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов: вибродиагностика, разработка метода вейвлет-анализа изделий машиностроения.

А.Н. Марков

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов: вычислительные системы, облачные вычисления (CLOUD COMPUTING), распределенные вычислительные системы, балансировка нагрузки вычислительных систем (load balancing).

А.Д. Стрельцов

Студент Белорусского государственного университета информатики и радиоэлектроники факультета компьютерного проектирования.

Аннотация. В статье рассматривается практика применения технологии *Big Data* предприятиями индустрии туризма, гостеприимства, авиаперевозок и других смежных отраслей, а также трудности, с которыми сталкиваются предприятия при внедрении *Big Data* в работу.

Под термином *Big Data* понимается технология, обеспечивающая решение задач по организации хранения, обработки, анализа и распределения больших объемов данных. Также в результате обработки и анализа данных технология *Big Data* помогает извлекать новые знания, которые в дальнейшем могут быть использованы для оптимизации бизнес-процессов предприятия таких как ценообразование, формирование привлекательного продукта и др.

Внедрение *Big Data* в работу предприятия предоставляет бизнесу ряд возможности и преимуществ. Однако на пути внедрения данной технологии в работу бизнес может столкнуться с рядом барьеров, создающих препятствие данному процессу. Такие барьеры включают в себя: недостаток квалифицированных кадров, обладающих необходимой подготовкой; большие затраты на внедрение; отсутствие достаточного опыта внедрения данной технологии в работу предприятий.

Ключевые слова: *Big Data*, туризм, информационные технологии, цифровая трансформация туризма, цифровая трансформация экономики.

Введение. Несмотря на ограничения, связанные с пандемией *COVID-19*, туризм по-прежнему остается модной нишей бизнеса. Здесь есть как известные крупные игроки, такие как *TripAdvisor*, *Booking.com*, *Airbnb*, так и малые и средние туристические агентства, которые ежедневно борются за обеспечение своей конкурентоспособности и лояльности

клиентов. Большие данные играют важную роль в достижении этих целей, помогая лучше понять потребности клиентов, проанализировать текущее состояние рынка и автоматизировать некоторые рабочие процессы.

В туризме, как и во многих других секторах бизнеса, большие данные отвечают за правильность принятия решений. Как правило, они занимаются прогнозированием потребительского спроса, персонализацией услуг, а также туристическим маркетингом и оптимизацией ценовой стратегии.

В целом большие данные обычно не являются самостоятельной технологией, поскольку требуют дополнительных методов их хранения, структурирования и анализа. Однако именно это помогает владельцам туристического бизнеса лучше понимать своих клиентов и предвидеть их потребности.

Определение больших данных. Понятие *Big Data* относится к наборам данных, объем которых не позволяет обрабатывать их, используя стандартные инструменты для работы с данными. Также *Big Data* можно определить, как набор технологий, предназначенных для получения информации. *Big Data* решает задачи хранения, обработки, анализа и распределения больших массивов данных, и при этом предоставляет возможности для работы с данными разной степени упорядоченности – это могут быть структурированные, неструктурированные и частично структурированные данные [1].

Внедрение технологии *Big Data* помогает получать новые знания о туристах в плане потребительского поведения: выбора направлений поездки, критериев выбора объектов размещения и др. Таким образом, на основе знаний, полученных в результате анализа больших объемов данных о поведении туристов, разрабатывается обобщенный портрет потребителя турпродукта, в котором учтены его предпочтения. В дальнейшем подобный образ путешественника может быть использован для формирования предложения, удовлетворяющего запросам рынка [2].

Областью применения технологии *Big Data* в туризме может быть информационное обеспечение продвижения туристских продуктов. В качестве примера может быть рассмотрен совместный проект *Profi.Travel* и *Amadeus* [3]. 13 сентября 2018 года *Profi.Travel* и *Amadeus* заключили соглашение о сотрудничестве и представили широкой аудитории новую технологию для продвижения туристических услуг – *Profi.Travel Programmatic*. *Profi.Travel Programmatic* дает возможность автоматизации процесса закупки рекламы с генерацией рекламного объявления, ориентированного на конкретного потребителя, на основе тревел-данных о пользователях (рисунок 1).

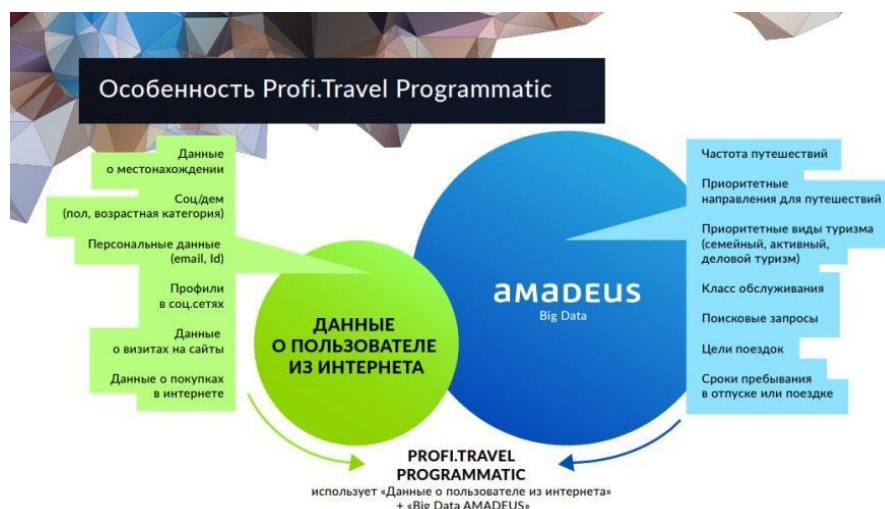


Рисунок 1. Архитектура Profi.Travel Programmatic

Другой успешный опыт применения технологии *Big Data* лежит в индустрии авиаперевозок – смежной по отношению к туризму отрасли. Так британская авиакомпания *British Airways* разработала и интегрировала в свою деятельность приложения и технологии, которые обеспечивают индивидуальный подход к путешественнику и понимание внутренних и внешних факторов, влияющих на решения всех стейкхолдеров. *British Airways* получает данные от клиента во время взаимодействия с ними в режиме реального времени, даже во время простой регистрации или в зонах ожидания. Вместе с тем *British Airways* использует результаты анализа накопленных данных о пассажирах-участниках программы лояльности авиакомпании в программе мониторинга потребительского спроса *Know Me*. Пример применения этой программы включает в себя отправку информации о конкретных клиентах в партнерские туристические агентства, которая может быть использована для формирования персонализированного предложения. Также, используя *Big Data*, *British Airways* в режиме реального времени получает и анализирует данные для поиска потерянного багажа, который отправляется в пункт назначения следующим рейсом, и не требует от пассажиров дополнительных действий для его поиска.

Другой успешной практикой внедрения технологии *Big Data* в деятельности авиакомпании является опыт швейцарского авиаперевозчика *Swiss International Airlines*, внедрившего технологии *Big Data* в процессы управления доходами и расходами компании. Процессы, включающие загрузку рейсов и ценообразования, были объединены в общий алгоритм с помощью цифровых инструментов для оперативного изменения цен.

Крупные отельные сети также используют технологии *Big Data* для анализа больших объемов данных в процессах управления доходами – анализе спроса и ценообразовании. Например, международная отельная сети *Marriott International* создала целые аналитические отделы, которые на основе анализа имеющихся данных о спросе и ценах на рынке формируют новое предложение. Другим примером применения технологии *Big Data* в индустрии гостеприимства является применение технологии «умного дома» в номерах отелей. Так система может автоматически регулировать работу кондиционера и электрического света в номере, чтобы добиться оптимальной температуры и яркости освещения. Таким образом, можно говорить о том, что технология *Big Data* расширяет возможности прямого контакта с клиентом и помогает достичь максимальной персонализации сервиса в различных условиях взаимодействия с ним [4].

Вклад больших данных в маркетинг туризма. Маркетинг, являющийся ключевой функцией туристического бизнеса, является идеальной основой для применения аналитики больших данных, что приводит к улучшению репутации, лояльности клиентов, финансовых показателей и позиции компании относительно конкурентов. В последнее десятилетие традиционные инструменты основывались на заранее определенном графике стратегии и поддерживали менталитет массового рынка, сосредотачивая внимание на человеке и его потребностях.

Туризм – это явление, характеризующееся изменчивостью, где планирование, программирование, спонтанность, риск, приключения и ожидания зависят от изменяющихся желаний туристов. Немедленная и целевая адаптация стратегий подхода приобретает благодаря большим данным персональный и уникальный характер.

Большие данные могут помочь понять, как думают клиенты и чего именно они хотят. Более того, благодаря инструментам прогнозирования можно прогнозировать будущие потребности, вызванные новыми популярными тенденциями в туризме.

Большие данные помогают собирать информацию о конкурентах и целевой аудитории на основе прошлогодних или даже более ранних отчетов, но в режиме реального времени. Никакая другая технология не даст такой же точности в исследованиях.

Учитывая разнообразие клиентов, будет сложно разработать единую целевую маркетинговую стратегию для их привлечения и удержания. В свою очередь, большие данные позволяют анализировать бесконечное количество представителей целевой аудитории, сегментировать их на группы и тем самым способствовать созданию таргетированного подхода к каждому типу.

В совокупности эту информацию можно использовать для точного прогнозирования и моделирования будущих бизнес-стратегий.

Данные о транзакциях (по операциям). Этот источник включает в себя данные веб-поиска, данные о посещениях веб-страниц, данные онлайн-бронирования и т. д. Обычно для получения этих данных используются расширенные веб-службы, такие как *Google Analytics*. (рисунок 2).

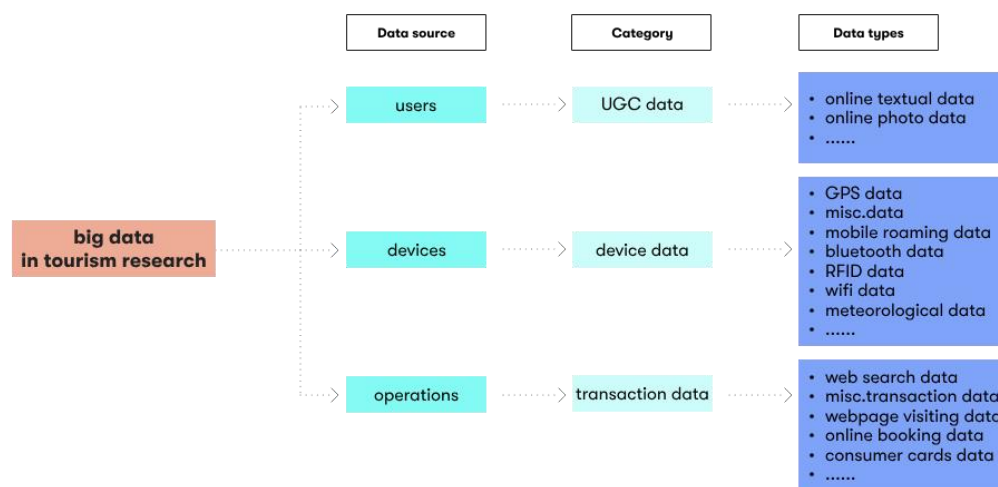


Рисунок 2. Данные о транзакциях

Большие данные, основанные на удовлетворенности существующих клиентов и ожиданиях будущих клиентов, помогают получить более точную информацию о потенциальной прибыли. В частности, процедура ценообразования упрощается за счет правильного прогнозирования пиковых периодов максимального спроса, а также актуальности услуг с учетом текущих тенденций рынка.

Текущие тенденции в области больших данных в туристической индустрии. В современном мире туристическая индустрия активно использует технологии больших данных для повышения прибыльности и улучшения обслуживания клиентов. Одной из ключевых тенденций в этой области является сбор и анализ огромных объемов данных, собранных из различных источников, таких как бронирования, отзывы туристов, социальные медиа, а также данные о местоположении и предпочтениях путешественников.

Один из способов использования больших данных в туризме заключается в анализе поведения клиентов и прогнозировании их потребностей. Путешественники оставляют ценную информацию в интернете, например, в виде отзывов о поездках, предпочтениях в отелях и местах, которые они посещают. Анализ этих данных позволяет туристическим компаниям понять предпочтения клиентов и адаптировать свои предложения под них.

Другим важным аспектом использования больших данных в туризме является оптимизация ценообразования и управление запасами. Анализ данных о бронированиях и спросе позволяет туристическим компаниям оптимизировать цены на свои услуги в реальном времени и управлять запасами, чтобы максимизировать прибыльность (рисунок 3).



Рисунок 3. Пример применения больших данных в туристической индустрии

Заключение. В результате проведенного анализа научной литературы и открытых источников, описывающих как технологию *Big Data* в целом, так конкретные примеры ее применения в индустрии туризма в частности, было установлено, что постоянно накапливающиеся большие объемы данных несут в себе потенциальные выгоды, которые можно извлечь, в том числе, и с помощью *Big Data*. Применение технологии *Big Data* помогает организовать хранение, передачу и обработку больших объемов данных, а также предоставляет методы для анализа и извлечения новых знаний, которые в дальнейшем могут быть использованы для получения значимых результатов, выраженных, применительно к индустрии туризма, в росте количества туристов, улучшении качества обслуживания, увеличении продаж и др.

Большие данные в туристической индустрии – это не просто трендовая технология, а инструмент для лучшего понимания ситуации на рынке и каждого клиента в целом. Всё это формирует отличную основу для более персонализированного подхода и более точного прогнозирования того, чего хотят клиенты.

Именно поэтому использование больших данных существенно усиливает конкурентное преимущество туристического бизнеса.

Список литературы

- [1] Величко Н.А. Технология Big Data. Анализ Рынка Big Data / Н.А. Величко, И.П. Митрейкин // Синергиянаук. 2018. № 30. С. 937-943.
- [2] Морозов М.М. Направления применения цифровых технологий в индустрии туризма и гостиничного бизнеса / М.М. Морозов // Стратегии и современные тренды развития предприятий туристского и гостиничного бизнеса: материалы IV Всероссийской научно-практической конференции (г. Москва, 23 апреля 2021 г.) –Москва, 2021. С. 352-358.
- [3] Имаметдинова М.Р. Ретроспективный анализ развития электронного бизнеса в сфере туризма / М.Р. Имаметдинова // Вестник РМАТ. 2020. № 1. С. 47-54.
- [4] Сергачева В.А. Использование Big Data в индустрии туризма / В.А. Сергачева, А.С. Додина, Г.Н. Кутепова // Индустрия туризма: возможности, приоритеты, проблемы и перспективы. 2019. Т. 14. № 1. С. 133-139.

Авторский вклад

Мигалевич Сергей Александрович – руководство исследованием потенциала современных информационных технологий в области туризма.

Марков Алексей Николаевич – рассмотрение текущих тенденций в области больших данных в туристической индустрии.

Стрельцов Андрей Дмитриевич – рассмотрение практики применения технологии *Big Data* предприятиями индустрии туризма, гостеприимства, авиаперевозок и других смежных отраслей.

BIG DATA PROCESSING IN THE TOURISM INDUSTRY

S.A. Migalevich
*Master of Technical Sciences,
Head of the Center for
Informatization and Innovative
Developments*

A.N. Markov
*Master of Technical Sciences,
Deputy Head of the Center for
Informatization and Innovative
Developments*

A.D. Streltsov
*Student of the Belarusian State
University of Informatics and
Radioelectronics*

Abstract. This article discusses practices of application of Big Data technology by enterprises, which operate in tourism, hospitality, airlines and other related areas, as well as challenges that they are facing during implication Big Data into their work. Based on these practices of application of Big Data technology and conducted survey of experts – employees of tourism industry enterprises – an analysis of practices of technology application in terms of challenges overcome and application areas is carried out.

The term Big Data is understood as a technology that provides a solution to the problems of organizing the storage, processing, analysis and distribution of large amounts of data. Also, as a result of data processing and analysis, Big Data technology helps to extract new knowledge, which can later be used to optimize the business processes of an enterprise, such as pricing, creating an attractive product, etc.

An application of Big Data technology to a work of an enterprise provides the business with a number of opportunities and benefits. However, on the way of applying of this technology to a work, a business may face a number of barriers that create an obstacle to this process. Such barriers include: lack of qualified personnel with the necessary skills; high implementation costs; lack of sufficient experience in applying of this technology to the work of enterprises.

Keywords: Big Data, tourism, information technology, digital transformation of tourism, digital transformation of economics.

УДК 004.056.53:004.75:004.67

ПРИМЕНЕНИЕ BIG DATA ДЛЯ ЗАЩИТЫ КОМПЬЮТЕРНЫХ СЕТЕЙ



М.В. Романюк

И.о. начальника отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики, магистрант кафедры ПИКС
romanuk@bsuir.by



Е.А. Лещенко

Инженер-программист отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики, магистрант кафедры ПИКС
e.leshchenko@bsuir.by



С.С. Марковский

Инженер-программист отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР
s.markovskij@bsuir.by

М.В. Романюк

Окончил Белорусский государственный университет информатики и радиоэлектроники. Инженер-программист, и.о. начальника отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики БГУИР, магистрант кафедры проектирования информационно-компьютерных систем БГУИР.

Е.А. Лещенко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Инженер-программист отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики БГУИР, магистрант кафедры проектирования информационно-компьютерных систем БГУИР.

С.С. Марковский

Окончил Белорусский государственный университет информатики и радиоэлектроники. Инженер-программист отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР.

Аннотация. Выполнен обзор применимости технологий Big Data для анализа нормального функционирования информационных систем и компьютерных сетей и защиты данных этих систем и сетей от нежелательных воздействий киберугроз.

Рассмотрен пример применения технологий *Big Data* в *SIEM*-системах и сделаны выводы о полезности применения *Big Data*, т.к. это позволяет эффективно анализировать миллиарды событий в день и не допускать возникновения инцидентов информационной безопасности.

Ключевые слова: защита компьютерных сетей, *SIEM*, прогнозирование угроз, информационная безопасность

Введение. В современном мире, где предприятия и организации в своей деятельности повсеместно используют компьютерные сети, обеспечение безопасности этих сетей имеет первостепенное значение. В связи с экспоненциальным ростом объема данных, генерируемых и передаваемых по этим сетям, традиционные меры безопасности стали недостаточными для борьбы с более сложными киберугрозами, которые появляются ежедневно.

Справляться с современными киберугрозами помогают технологии *Big Data*, предлагающие расширенные возможности обработки и анализа больших объемов данных для обнаружения и устранения угроз безопасности в режиме реального времени. В этой статье рассматривается применение больших данных для повышения безопасности компьютерных сетей, их преимущества, проблемы и будущие перспективы.

Сбор и агрегирование данных. Аналитика *Big Data* начинается со сбора и агрегирования данных из различных источников в сети. Сюда входят сетевые журналы, пакеты, системные события и журналы действий пользователей. Собирая данные из различных источников, платформы *Big Data* создают комплексное представление о деятельности сети, позволяя специалистам информационной безопасности выявлять нехарактерные паттерны и потенциальные угрозы.

Мониторинг и обнаружение в режиме реального времени. Одним из ключевых преимуществ Больших Данных в области сетевой безопасности является возможность мониторинга и анализа сетевого трафика в режиме реального времени. Передовые алгоритмы аналитики могут обнаруживать аномалии, такие как необычные скачки трафика, попытки несанкционированного доступа или подозрительные модели поведения, немедленно информируя специалистов информационной безопасности о потенциальных угрозах. Такой подход позволяет быстро реагировать на угрозы и устранять их последствия, сводя к минимуму влияние инцидентов безопасности [1].

Поведенческий анализ и обнаружение аномалий. Аналитика больших данных позволяет применять методы машинного обучения и искусственного интеллекта для поведенческого анализа и обнаружения аномалий. Создавая базовые профили поведения пользователей, устройств и приложений эти системы могут выявлять отклонения от нормальных моделей, свидетельствующие о нарушениях безопасности или внутренних угрозах. Такой подход к обеспечению безопасности позволяет организациям обнаруживать угрозы и реагировать на них до того, как они перерастут в серьезные инциденты информационной безопасности.

Прогнозирующая аналитика и анализ угроз. Платформы *Big Data* могут использовать исторические данные и информацию об угрозах для прогнозирования и предотвращения будущих инцидентов безопасности. Анализируя прошлые атаки и тенденции, эти системы могут предвидеть возникающие угрозы и уязвимости, позволяя организациям принимать упреждающие меры для обеспечения безопасности. Кроме того, интеграция внешних источников информации об угрозах обеспечивает ценный контекст для понимания развивающегося спектра угроз, позволяя специалистам информационной безопасности действовать на опережение злоумышленников.

Масштабируемость и гибкость. Технологии *Big Data* обеспечивают масштабируемость и гибкость, позволяя организациям адаптироваться к динамичному характеру сетевых сред. Будь то крупномасштабные распределенные сети или облачные инфраструктуры, платформы *Big Data* могут легко справиться с объемом, скоростью и разнообразием генерируемых данных, обеспечивая непрерывный мониторинг и анализ без снижения производительности.

Примеры применения *Big Data* в защите компьютерных сетей. Хорошим примером использования технологий *Big Data* для повышения безопасности в компьютерных сетях предприятий и организаций являются *SIEM*-системы.

Security Information and Event Management (SIEM) – это программные комплексы для обеспечения компьютерной безопасности, которые сочетают в себе мониторинг ИТ-инфраструктуры в режиме реального времени на предмет угроз безопасности со сбором и анализом данных журналов и событий с различных компонентов *IT*-среды.

Появление больших объемов быстро меняющихся неструктурированных данных, таких как данные, генерируемые веб-приложениями, электронной почтой и социальными

сетями, создает проблему для SIEM с точки зрения способности находить корреляции и другую важную информацию о текущем состоянии безопасности. Такая информация может быть важна для обеспечения информационной безопасности, но старые типы SIEM-систем могут быть не приспособлены для работы с ней.

Сила технологии больших данных такова, что теперь доступны распределенные вычислительные среды и фреймворки, такие как *Hadoop*, позволяющие легко хранить и анализировать огромные объемы неструктурированных данных. Аналитика больших данных позволяет расширить возможности SIEM-систем в обнаружении угроз, предоставляя доступ к корреляции между пулами данных, которая ранее была недоступна. Эти большие данные могут включать в себя файлы журналов и события из внутренних систем, а также внешние источники, такие как данные разведки угроз, базы данных уязвимостей и данные социальных обогатителей [2].

Объем данных о событиях безопасности настолько велик, что специалисты информационной безопасности крупнейших современных предприятий могут анализировать миллиарды событий в день. Решения SIEM, интегрированные с инфраструктурой больших данных, позволяют организациям избежать сложностей, связанных с обработкой всех данных и событий, необходимых для обнаружения угроз информационной безопасности, при отсутствии в организации сложной инфраструктуры и больших человеческих ресурсов для анализа всей информации и поиска таких нарушений.

Заключение. Таким образом, применение *Big Data* для обеспечения информационной безопасности компьютерных сетей представляет собой значительный сдвиг в парадигме обнаружения и устранения угроз. Используя возможности передовой аналитики, машинного обучения и мониторинга в режиме реального времени, технологии *Big Data* позволяют специалистам информационной безопасности опережать развивающиеся киберугрозы и защищать критически важные активы. Однако эффективность решений по обеспечению безопасности на основе *Big Data* зависит не только от технологий, но и от квалифицированного персонала и надежных стратегий кибербезопасности. По мере того, как организации будут продолжать цифровые преобразования, интеграция аналитики *Big Data* в их системы безопасности будет иметь решающее значение для поддержания устойчивости к постоянно меняющемуся спектру угроз.

Список литературы

[1] Leveraging Big Data Analytics with SIEM: Extracting Actionable Insights for Proactive Security: [Электронный ресурс]. URL: <https://www.linkedin.com/pulse/leveraging-big-data-analytics-siem-extracting-aravind-raghunathan>. (Дата обращения: 09.02.2024).

[2] What is SIEM and How Does it Relate to Big Data and Machine Learning? [Электронный ресурс]. URL: <https://datafloq.com/read/siem-how-does-relate-big-data-machine-learning/>. (Дата обращения: 11.02.2024).

Авторский вклад

Авторы внесли равноценный вклад.

APPLYING BIG DATA TO PROTECT COMPUTER NETWORKS

M.V. Romaniuk

Acting Head of the Network Technologies Department of the Center of Informatization and Innovation Elaborations, Assistant of the Department of Computer Science of BSUIR, Master's student of the Department of ICSD of BSUIR

E.A. Leshchenko

Software Engineer of the Network Technology Department of the Center of Informatization and Innovation Elaborations, Assistant of the Department of Computer Science of BSUIR, Master's student of the Department of ICSD of BSUIR

S.S. Markovskii

Software Engineer of the Network Technology Department of the Center of Informatization and Innovation Elaborations

Abstract. The review of the applicability of Big Data technologies for analyzing the normal functioning of information systems and computer networks and protecting the data of these systems and networks from the undesirable effects of cyber threats is carried out.

An example of the use of Big Data technologies in SIEM systems is considered and conclusions are drawn about the usefulness of using Big Data, since it allows you to effectively analyze billions of events per day and prevent the occurrence of information security incidents.

Keywords: protection of computer networks, SIEM, threat prediction, information security

УДК 005.511:004.42

ИСПОЛНЯЕМОЕ МОДЕЛИРОВАНИЕ БИЗНЕС-ПРОЦЕССОВ В СИСТЕМЕ BIZAGI PROCESS MODELER



Е.А. Бущик
аспирант кафедры
информационных
радиотехнологий БГУИР,
e.bushchik@bsuir.by

Е.А. Бущик

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с цифровизацией процессов управления в системе среднего специального учреждения образования, использованием цифровых технологий для оптимизации бизнес-процессов.

Аннотация. В статье рассматривается моделирование бизнес-процесса с применением соответствующего специализированного программного обеспечения. Проанализированы популярные инструментальные средства моделирования бизнес-процессов в нотации BPMN, в частности, на примере описания бизнес-процесса «Планирование образовательного процесса» демонстрируется исполняемое моделирование при помощи инструментальных средств в системе Bizagi Process Modeler.

Ключевые слова: бизнес-процесс, исполняемое моделирование, автоматизация бизнес-процессов, инструментальные средства моделирования бизнес-процессов, нотация.

Введение. Вектор развития современного общества находится под воздействием цифровых технологий. Изменения касаются не только различных отраслей промышленности, где роботизация является приметой настоящего времени, но стремительная цифровизация внесла изменения и в систему государственного управления, научную, культурную и образовательную сферы. Перед системой образования процесс цифровизации ставит серьезный вызов, делая акцент на опережающую подготовку выпускников учреждений образования, владеющими компетенциями, востребованными в будущем.

В соответствии с Концепцией развития системы образования Республики Беларусь до 2030 года, в качестве основных направлений развития определены: совершенствование организации образования, его содержания, научно-методического обеспечения и процессов управления [1]. В рамках направления совершенствования процессов управления для осуществления цифровизации необходимо провести их оптимизацию. Оптимизация бизнес-процессов означает разработку и реализацию методов по изменению бизнес-процессов, чтобы выявить наиболее оптимальный путь для достижения поставленных целей.

Начинать деятельность по оптимизации требуется с описания существующих бизнес-процессов, которое предполагает создание моделей бизнес-процессов для наглядного отображения процесса с помощью составляющих: данные, действия, события.

Существует три основных подхода к бизнес-моделированию:

- функциональное моделирование – это вид моделирования, который подразумевает описание процессов в виде взаимосвязанных, четко структурированных функций;
- процессное моделирование – описание деятельности предприятия в виде бизнес-процессов, детализируемых в непрерывные взаимосвязанные функции;
- ментальное моделирование – описание процесса как набора связанных друг с другом понятий.

Для бизнес-моделирования определен процессный подход, так как он позволяет сосредоточиться не только на отдельных функциях или задачах, но и на эффективном выполнении взаимосвязанных и взаимозависимых бизнес-процессов. При реализации процессного моделирования могут быть применены следующие нотации: *EPC*, *BPMN*, *IDEF3*, *VAD*, *SIPOC*, *Flow Charting*.

Анализ систем моделирования в нотации *BPMN*. При проектировании в рамках исследования бизнес-процессов учреждения среднего специального учреждения образования была выбрана нотация *BPMN* по ряду следующих достоинств:

- большой набор графических элементов, объединенных по своим функциональным назначениям в типы моделей;
- подробное и разностороннее описание процедур и операций;
- удобные взаимосвязи между пересекающимися действиями;
- выраженная гибкость нотации;
- возможность использования различными аудиториями.

Авторы *BPMN* выделяют три уровня моделирования:

- 1) описательный – простое моделирование, описывающее процесс на бизнес-уровне с использованием базовых элементов нотации;
- 2) аналитический – моделирование, включающее альтернативные пути и исключения, требующее описывать процесс с максимально необходимой детализацией;
- 3) исполняемый – моделирование процесса, пригодного для исполнения в *BPMS/BPMT* (*Business Process Management System/Tool*) системах.

Исполняемое моделирование предполагает использования исполняемой среды. В качестве среды исполнения бизнес-процессов можно использовать систему управления бизнес-процессами *BPMS/BPMT* или процессно-ориентированную информационную систему, которая поддерживает бизнес-процессы. На данный момент наиболее характерными и распространенными программными продуктами, реализующими исполняемое моделирование бизнес-процессов являются:

- *Microsoft Visio*;
- *Enterprise architect*;
- *ELMA BPM*;
- *Bizagi Process Modeler*;
- *ARIS Express*;
- *Modelio*.

Для выбора одного из вышеперечисленных программных продуктов проводилась сравнительная характеристика по следующим критериям: язык интерфейса, стоимость, *no-code BPM* система, визуальный редактор, режим отладки. Результаты сравнения представлены в таблице 1.

Таблица 1. Сравнительная характеристика ПО для моделирования бизнес-процессов

№	Наименование	Язык интерфейса	Стоимость	No-code BPM система	Визуальный редактор	Режим отладки
1	<i>Microsoft Visio</i>	Русский	Лицензионный ключ (бессрочный) – 1500 рос.руб	да	да	да
2	<i>Enterprise architect</i>	<i>English</i>	Виды лицензий: – <i>professional: standart and floating</i> – 299\$; – <i>corporate: standart</i> – 299\$, <i>floating</i> – 399\$; – <i>unified: standart</i> – 499\$, <i>floating</i> – 649\$; – <i>ultimate: standart</i> – 699\$, <i>floating</i> – 899\$.	да	да	да
3	<i>ELMA BPM</i>	Русский	Лицензия ELMA365 Standard в облачном сервисе SaaS (стоимость хостинга) – 72 \$/год; с размещением на собственном сервере On-Premises – 144 \$/год на одного пользователя.	да	да	да
4	<i>Bizagi Process Modeler</i>	Русский, <i>English</i>	Лицензия бесплатная	да	да	да
5	<i>ARIS Express</i>	<i>English</i>	Лицензия бесплатная	да	да	да
6	<i>Modelio</i>	<i>English</i>	Лицензия бесплатная	да	да	да

В качестве системы *BPMS/BPMT* была выбрана *Bizagi Process Modeler* – это *BPM*-система, разработанная одноименной компанией, которая предлагает использовать платформу для быстрой автоматизации процессов любого рода через механизмы «*drag-and-drop*», способную учитывать персонализированный контекстный опыт разработчиков [2].

Пример моделирования бизнес-процесса в *Bizagi Process Modeler*. Рассмотрим моделирование бизнес-процесса в системе *Bizagi Process Modeler* на примере бизнес-процесса «Планирование образовательного процесса». Построение бизнес-модели происходит под

контролем системы, например, встроенные правила соединения элементов гарантируют соответствие нотации *BPMN* и формируют подсказки, направляющие работу пользователя как показано на рисунке 1.

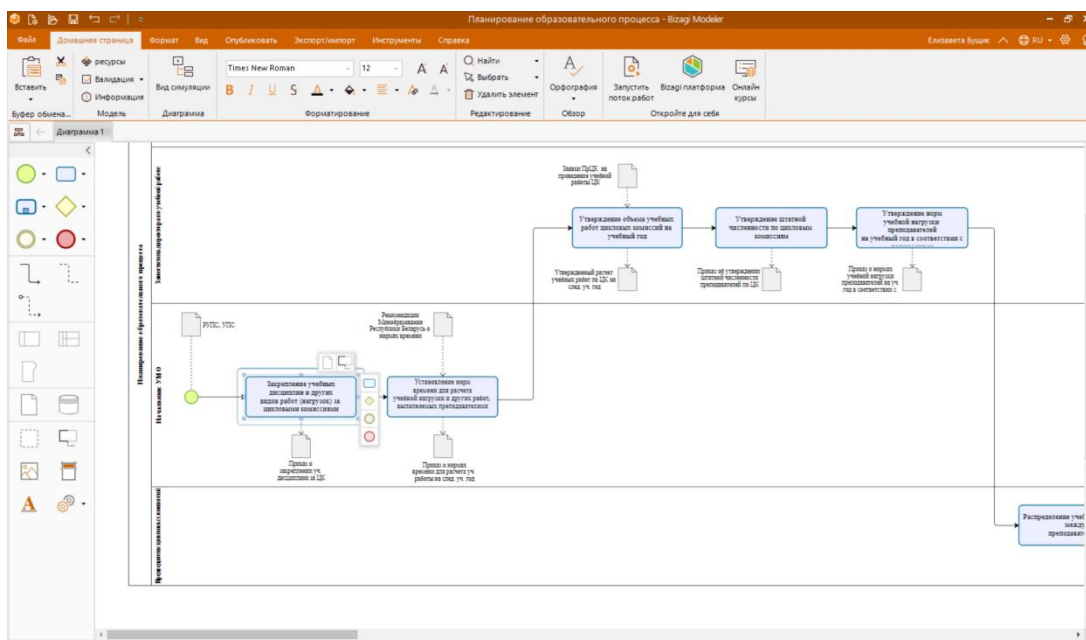


Рисунок 1. Модель бизнес-процесса в Bizagi Process Modeler

Разработка модели бизнес-процесса выполняется с помощью панелей элементов, соответствующих элементам нотации *BPMN*. Большую наглядность моделей в *Bizagi* обеспечивает применение пулов, служащих для логической группировки операций процесса и дорожек, визуализирующих исполнителей. Симуляция бизнес-процессов в *Bizagi* выполняется в разрезе исполнителей, при этом доступен мониторинг порядка выполнения задач.

Действия бизнес-процесса представляются с помощью карты маршрута, которая, как видно на рисунке 2, описывает логику выполнения бизнес-процесса от начальной точки до точки завершения.

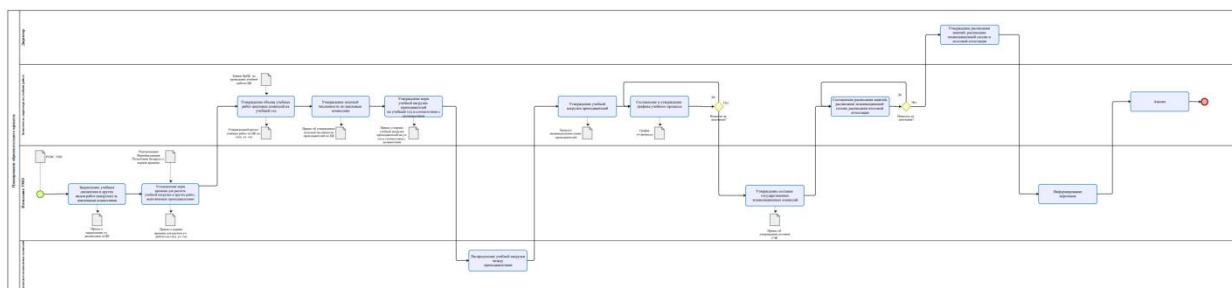


Рисунок 2. Бизнес-процесс «Планирование образовательного процесса» в нотации *BPMN*

Карта маршрута составляется из элементов, являющихся подмножеством *BPMN*. Задачи в модели соответствуют пользовательским заданиям. Выполнение задачи приводит к продвижению бизнес-процесса от одной точки к другой.

Заключение. В статье рассмотрены возможные подходы к моделированию исполняемых бизнес-процессов нотации *BPMN*. Моделирование выполнено при помощи средств программного обеспечения *Bizagi Process Modeler*, представляющих собой процессно-ориентированную систему

автоматизации деятельности организации. Таким образом, посредством моделирования предоставляются возможности:

- стандартизировать работу сотрудников;
- проанализировать потенциальные возможности усовершенствования;
- определить проблемы в целях их устранения;
- снизить расходы без потери качества;
- определить требования к человеческим, информационным, финансовым ресурсам, обеспечивающим эффективное выполнение операций.

Моделирование бизнес-процессов в дальнейшем должно обеспечить создание унифицированной информационной системы учреждения образования, что в свою очередь откроет возможность для реализации единой образовательной среды учреждений образования среднего специального уровня образования.

Список литературы

- [1] Концепция развития системы образования Республики Беларусь до 2030 года [Электронный ресурс]. – Режим доступа: <https://pravo.by/document/?guid=3871&p0=C22100683>.
- [2] Bizagi Studio [Электронный ресурс]. Режим доступа: <https://www.bizagi.com/platform/studio>.

Авторский вклад

Бущик Елизавета Александровна – анализ программного обеспечения исполняемого моделирования в нотации *BPMN*, моделирование бизнес-процесса «Планирование образовательного процесса» в *BPM*-системе *Bizagi Process Modeler*.

EXECUTABLE SIMULATION BUSINESS PROCESSES IN BIZAGI PROCESS MODELER

L.A. Bushchik

*Postgraduate student of the
Department of Radio Information
Technologies of BSUIR*

Abstract. The article deals with business process modeling with the use of appropriate specialized software. Popular tools for modeling business processes in *BPMN* notation are analyzed, in particular, on the example of business process description «Planning of educational process» the executable modeling with the help of tools in *Bizagi Process Modeler* system is demonstrated.

Keywords: business process, executable modeling, business process automation, business process modeling tools, notation.

Научное издание

BIG DATA И АНАЛИЗ ВЫСОКОГО УРОВНЯ

BIG DATA AND ADVANCED ANALYTICS

**СБОРНИК НАУЧНЫХ СТАТЕЙ
X МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ
КОНФЕРЕНЦИИ**

(Республика Беларусь, Минск, 13 марта 2024 года)

В двух частях
Часть 2

В авторской редакции
Ответственный за выпуск *Д. В. Лихачевский*
Компьютерная верстка *Е. М. Косарева, И. А. Евдокимова*

Подписано в печать 27.02.2024. Формат 60×84 1/8. Бумага офсетная. Гарнитура «Таймс».
Отпечатано на ризографе. Усл. печ. л. 57,89. Уч.-изд. л. 49,3. Тираж 25 экз. Заказ 16.

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



ISBN 978-985-543-752-0



9 789855 437520